



The Effectiveness of a School-Based Test Anxiety Intervention at Primary-Level

Alison Mc Fadden

Supervisors: Dr. Fionnuala Tynan and Dr. John Perry

Thesis submitted in partial fulfillment of requirements for the award of Doctorate in Educational and Child Psychology

Submitted to Mary Immaculate College, August 2022

Abstract

Background

Test anxiety (TA) is characterised as an extreme fear of assessment which has a debilitating effect on performance. Despite growing concern over TA in primary-level education in Ireland, research on evidence-based intervention is limited.

Aims

The aim of this research was to determine the effectiveness of a school-based TA intervention in an Irish primary setting. The goal of this was to extend the research on TA intervention at primary-level, and subsequently, inform Educational Psychologists of evidence-based intervention for practice. Weems' (2015) programme was chosen for evaluation as it has shown promising evidence in reducing TA in targeted group settings. In the move towards inclusive educational practice, this study aimed to determine the effectiveness of the programme in a universal setting.

Methods

Prior to the main study, a systematic review was carried out to examine the existing research on school-based TA intervention. Eight studies were identified in this review. It was concluded that the evidence for school-based intervention remains limited due to methodological weaknesses, statistically non-significant findings, and a lack of generalisability to an Irish context. This review revealed promising evidence for Weems' (2015) multi-modal TA programme. Consequently, this programme was chosen for evaluation in the main study.

The main study employed a quantitative research design in the form of a cluster randomised control trial. Two fourth class groups were randomly assigned to an intervention group ($n = 22$) or a waitlist control group ($n = 17$). TA was measured as the primary outcome with 'self-efficacy for academic achievement' and 'emotion regulation' measured as secondary outcomes. Two emotion regulation strategies were examined, 'cognitive reappraisal' (CR) and 'expressive suppression' (ES). Data was compared between groups and within groups at three timepoints: pre-intervention, post-intervention, and six-week follow-up.

Results

Data analysis of participants included in pre-intervention and post-intervention data revealed a statistically significant reduction in TA in the intervention group from pre-intervention to post-intervention with a moderate effect size, whilst the control group did not demonstrate any significant changes. There was no statistically significant time*group interaction, however, despite a medium effect size. TA outcomes were maintained at a six-week follow-up with a large effect size from pre-intervention, and with further significant reductions from post-intervention to six-week follow-up with a small effect size in the follow-up sample of the intervention group. Significant reductions were also observed from pre-intervention to post-intervention and to follow-up in the control group follow-up sample, although a high attrition rate threatened the validity of these findings. A significant reduction in the use of ES strategies was also observed in the intervention group with a moderate effect size in a 2x2 ANOVA, although this significance was not detected in a 3x2 ANOVA; this was attributed to the increased comparisons of this analysis which reduces statistical power. No significant improvements were observed in academic self-efficacy or CR across either group or at any timepoint.

Conclusions

The findings from this initial study provide promising evidence for the effectiveness of this intervention as a universal programme in reducing TA and ES in Irish primary-level pupils, although preliminary given the absence of time*group interactions, mixed findings for the original participant sample and follow-up sample, and study limitations. Future research which addresses the limitations of this study, for example with a larger sample size and broader age group, would be beneficial in strengthening this evidence and generalisability of results.

Key words: Test anxiety, universal intervention, inclusive education, wellbeing

Declaration

I, Alison Mc Fadden, hereby declare that the work presented in this thesis is my own. Where information has been derived from other sources, this has been acknowledged and referenced.

Alison Mc Fadden

Alison Mc Fadden

Date: 24/08/2022

Word count (excluding tables and figures): 33,471

Acknowledgements

I would like to offer my sincere gratitude to the school who took part in this study. I am extremely grateful for the hard work and dedication of all pupils involved. I would also like to thank the teachers who generously offered their class time for this research, as well as the parents who volunteered their time to complete questionnaires.

To my research supervisors, Dr. Fionnuala Tynan and Dr. John Perry, thank you for your support and guidance throughout this research process. I have valued your insight and expertise in refining this project from start to finish. Finally, your enthusiasm, reassurance, and positivity were a continuous source of encouragement for me in completing this research.

To my professional placement supervisor who supported me in my quest to source an appropriate school for this research, I am truly grateful.

To my DECPsy peers, I could not have asked for a better support system during this doctoral training. Best of luck to everyone in your future careers.

To my family, your unwavering support and words of encouragement throughout my studies have meant the world to me. I could not have done it without you.

To my friends, thank you for your support and understanding every step of the way.

Finally, to Dean, I appreciate every moment you have been there for me, offering love, patience, and a constant source of laughter.

Table of Contents

Chapter One: Introduction	1
1.1. Key Issue- Test Anxiety	1
1.2. Focus and Purpose of Thesis	1
1.3. Personal Interest	2
1.4. Research Questions	3
1.4.1. Primary Research Question	3
1.4.2. Secondary Research Questions.....	3
1.5. Thesis Structure.....	3
Chapter Two: Review Paper	3
Chapter Three: Empirical Paper	4
Chapter Four: Critical Review and Impact Statement.....	4
Chapter Two: Review Paper	5
2.1. Introduction to Anxiety	5
2.1.1. Definition.....	5
2.1.2. Origins of Anxiety	5
2.1.3. Childhood Anxiety	6
2.2. Overview of Test Anxiety	7
2.2.1. Definitions and Theories of Test Anxiety	7
2.2.2. Manifestations of Test Anxiety	12
2.2.3. Test Anxiety: Contributing and Correlational Factors	14
2.3. The Issue of Test Anxiety	16
2.3.1. Increasing Prevalence of Test Anxiety at Primary-level	17
2.3.2. Consequences of Test Anxiety	18
2.4. Focus and Purpose of Review in an Educational Psychology Practice and Policy Context	19
2.5. Existing Intervention Research	20
2.6. Systematic Review of the Literature	21
2.6.1. Review Question.....	21
2.6.2. Search Strategy	21
2.6.3. Weighting of Studies Framework.....	26
2.6.4. Participants	27
2.6.5. Interventions	28
2.6.6. Measures	29

2.6.7. Study design	30
2.6.8. Summary of Intervention Outcomes.....	31
2.6.9. Synthesis of Findings.....	34
2.6.10. Outstanding Issues and Future Research	37
Chapter Three: Empirical Paper	41
3.1. Introduction	41
3.1.1. Key Issue- Test Anxiety at Primary-Level.....	41
3.1.2. Existing Intervention Research.....	41
3.1.3. Aims of the Present Study	42
3.2. Research Questions and Hypotheses.....	42
3.2.1. Primary Research Question and Hypotheses.....	42
3.2.2. Secondary Research Questions and Hypotheses	43
3.3. Method	43
3.3.1. Research Paradigm	43
3.3.2. Research Design and Procedure	43
3.3.3. Participant Sample.....	44
3.3.4. Ethical Considerations	46
3.3.5. Intervention.....	47
3.3.6. Measures.....	51
3.3.7. Data Analysis.....	53
3.4. Results	54
3.4.1. Comparison of Baseline Data Between Groups	54
3.4.2. Skewness and Kurtosis Analyses	55
3.4.3. Homogeneity of Variance.....	56
3.4.4. Intervention Effects on Test Anxiety.....	56
3.4.5. Intervention Effects on Academic Self-Efficacy.....	59
3.4.6. Intervention Effects on Emotion Regulation	60
3.4.7. Additional Observations	62
3.5. Discussion	63
3.5.1. Key Findings and Implications.....	64
3.5.2. Strengths and Limitations of the Present Study.....	71
3.5.3. Additional Recommendations for Future Research.....	77
3.6. Conclusion.....	78

Chapter Four: Critical Review and Impact Statement	79
4.1. Epistemological and Ontological Perspective.....	79
4.2. Strengths & Limitations of the Study.....	80
4.2.1. Research Design.....	81
4.2.2. Sample.....	81
4.2.3. Measures.....	82
4.2.4. Intervention.....	84
4.2.5. Universal Delivery.....	85
4.2.6. Timing of the Intervention.....	86
4.2.7. Hawthorne Effect.....	87
4.3. Implications of the Research.....	88
4.3.1. Implications for Knowledge of Universal TA Intervention.....	88
4.3.2. Implications for Educational Policy and Curriculum.....	89
4.3.3. Implications for Practice.....	91
4.3.4. Implications for Future Research.....	93
4.4. Distinct Contribution.....	94
4.5. Personal Reflection.....	95
4.5.1. Description.....	95
4.5.2. Feelings.....	95
4.5.3. Evaluation.....	96
4.5.4. Analysis.....	97
4.5.5. Conclusion.....	98
4.5.6. Action Plan.....	98
4.6. Impact Statement.....	98
4.7. Dissemination of Research.....	100
References	101
Appendices	126

List of Tables

Table 1:	Symptoms of Test Anxiety (Kennedy, 2010; Sawka-Miller, 2011).....	13
Table 2:	Search Terms for Online Databases.....	22
Table 3:	Inclusion and Exclusion Criteria.....	23
Table 4:	Overall WoE Scores.....	26
Table 5:	Demographic Information Across Participant Groups.....	45
Table 6:	Overview of Intervention Sessions.....	47
Table 7:	Pairwise Comparisons of Pre-Intervention Data Between Groups for Original Sample.....	54
Table 8:	Pairwise Comparisons of Pre-Intervention Data Between Groups for Follow-up Sample.....	55
Table 9:	Summary of Skewness and Kurtosis Scores.....	55
Table 10:	P-Values for Levene’s Homogeneity of Variance.....	56
Table 11:	Summary of Means and Standard Deviations for Pre-Intervention and Post Intervention Test Anxiety Scores on the Children’s Test Anxiety Scale (CTAS).....	57
Table 12:	Summary of Means and Standard Deviations for Pre-Intervention, Post Intervention and Follow-Up Test Anxiety Scores on the Children’s Test Anxiety Scale (CTAS) for Follow-Up Sample.....	59
Table 13:	Summary of Means and Standard Deviations for Baseline, Post- Intervention and Follow-up Scores on the Children’s Self-Efficacy Scale for Academic Achievement.....	60
Table 14:	Summary of Means and Standard Deviations for Baseline, Post- Intervention and Follow-up Cognitive Reappraisal Scores on the Emotion Regulation Questionnaire for Children and Adolescents.....	61
Table 15:	Summary of Means and Standard Deviations for Baseline, Post- Intervention and Follow-up Expressive Suppression Scores on the Emotion Regulation Questionnaire for Children and Adolescents.....	62

List of Figures

Figure 1:	Yerkes-Dodson (1908) Law as cited in Kahneman (1973, p.34).....	8
Figure 2:	Flow Chart of the Final Search (2021) Screening Process for Article Selection.....	25
Figure 3:	NEPS Continuum of Support within the Wellbeing Policy Statement and Framework for Practice.....	39
Figure 4:	NEPS Continuum of Support within the Wellbeing Policy Statement and Framework for Practice.....	92

List of Appendices

Appendix A: List of Excluded Studies

Appendix B: Final List of Studies Included in Review

Appendix C: Summary of Included Studies

Appendix D: Weight of Evidence

Appendix E: Example of Kratochwill's (2003) Coding Protocol Completed for One Study

Appendix F: Invitation Letter to School Principal

Appendix G: Teacher Information Sheet

Appendix H: Teacher Consent Form

Appendix I: Parent Information Sheet

Appendix J: Parent Consent Form

Appendix K: Participant Information Sheet

Appendix L: Participant Assent Form

Appendix M: Ethical Approval Forms

Appendix N: Weems' (2015) Programme Manual

Appendix O: Programme Fidelity Checklist

Appendix P: Children's Test Anxiety Scale

Appendix Q: Adapted Children's Self Efficacy Scale- Academic Achievement

Appendix R: Emotion Regulation Questionnaire for Children and Adolescents

Appendix S: Study Limitations and Considerations for Future Research

List of Abbreviations

ANOVA	Analysis of Variance
APA	American Psychiatric Association
CBT	Cognitive Behavioural Therapy
CPD	Continued Professional Development
CR	Cognitive Reappraisal
CSES-AA	Children's Self-Efficacy Scale- Academic Achievement
CoS	Continuum of Support
CTAS	Children's Test Anxiety Scale
DES	Department of Education and Skills
DSM-5	Diagnostic and Statistical Manual of Mental Health Disorders- Fifth Edition
EP	Educational Psychologist
ES	Expressive Suppression
ERQ-CA	Emotion Regulation Questionnaire for Children and Adolescents
MIREC	Mary Immaculate Research Ethics Committee
NCCA	National Council for Curriculum and Assessment
NEPS	National Educational Psychological Service
PE	Physical Education
PSI	Psychological Society of Ireland
RCT	Randomised Control Trial
SPHE	Social, Personal, and Health Education
TA	Test Anxiety
TEP	Trainee Educational Psychologist
WoE	Weight of Evidence

Chapter One: Introduction

Chapter one of this thesis introduces the key concepts and relevant issues of this initial study. To begin with, a definition of the core concept of test anxiety (TA) is presented, followed by the issue of TA in an Irish primary-level context. Next, the focus and purpose of the thesis, as well as the researcher's personal interest in the topic, are discussed. Finally, a brief outline of the remaining structure of the thesis is provided.

1.1. Key Issue- Test Anxiety

Testing and evaluative situations have emerged as a pervasive class of stressors in Western society (Zeidner, 2007). Although stress and nervousness are considered normal reactions to evaluative settings, when feelings of anxiety become so extreme or intense that they interfere with performance, it is classed as TA (Larsen, 2017). TA is a universal phenomenon which is commonly experienced by students of all ages (Columbus, 2008). In recent years, TA has become increasingly prevalent in primary-level children and is believed to be attributed to a growing emphasis on standardised testing (Lobman, 2014). In an Irish context, it has been reported that 50-75% of primary-level teachers feel that standardised testing causes extreme anxiety in pupils (Devine et al., 2020; O'Leary et al., 2019). In line with a national emphasis on wellbeing (Department of Education and Skills [DES], 2018, 2022b), these reports highlight the need for evidence-based proactive interventions in primary-level settings.

1.2. Focus and Purpose of Thesis

The primary focus of this thesis is the evaluation of a TA intervention (Weems, 2015) delivered universally in an Irish primary-level context. School-based interventions are essential in addressing the issue of TA (Yew Chye, 2008), as the educational setting is where assessment situations and subsequent TA arise (Yeo et al., 2016). The purpose of this initial research is to extend the literature on school-based TA intervention at primary-level and to determine the effectiveness of this programme in an Irish context. In turn, it is hoped that extending the research will inform Educational Psychologists (EPs) of effective intervention in their role of 'scientist practitioners' (Birch et al., 2015). In this role, EPs have a responsibility to seek the best available evidence to inform effective psychological services (Lane & Corrie, 2007). The goal of evidence-based practice is to optimise positive outcomes for the individuals receiving intervention (Davidson, 2005). Previous review studies have concluded that research on TA intervention at primary-level is extremely limited (Ergene,

2003; Von Der Embse et al., 2013), despite growing concern over TA at this level (Lobman, 2014; O'Leary et al., 2019). Weems' (2015) school-based TA intervention was chosen for evaluation as it has shown promising evidence among populations of children and adolescents in the United States of America (USA) in targeted groups settings (Weems et al., 2014; Weems et al., 2009). In consideration of the modern-day emphasis on inclusive educational practice (Dawson & Guare, 2018), this research aimed to determine the effectiveness of the intervention as a universal programme. This proactive approach is essential in line with the Wellbeing Policy Statement and Framework for Practice (DES, 2018). This is a national policy outlining the role of educators in preparing children and young people for those life challenges which may influence their wellbeing.

A secondary aim of this thesis was to evaluate intervention effects on academic self-efficacy as this is considered as a causal and consequent factor of TA. It is also targeted for improvement in Weems' (2015) TA intervention. Academic self-efficacy refers to the subjective judgement which a person makes regarding their own academic capabilities (Bandura, 1986).

Finally, as the selected intervention targets emotion management with the aim of supporting pupils to cope with experiences of TA, an additional secondary aim was to determine changes in emotion regulation skills following programme completion. Specifically, two types of emotion regulation skills were examined, 'cognitive reappraisal' (CR) and 'expressive suppression' (ES). CR refers to a process of reframing an emotional stimulus to reduce its emotional impact (Troy et al., 2018). Alternatively, ES is the attempt to conceal emotional reactions from others in social interactions (Butler et al., 2003).

1.3. Personal Interest

My interest in the topic of TA and relevant interventions was first sparked during a lecture on maths assessment and intervention. This lecture brought about a discussion on maths anxiety. I was intrigued by this notion of maths anxiety and began to search the literature to learn more. From my reading, I discovered that maths anxiety often occurs in the context of testing situations (Gilmore et al., 2018), and is conceptually related to TA given the common fear of evaluation and impact on performance (Zeidner & Matthews, 2005). I came to realise that TA was a broader and more prevalent issue, and often accounted for experiences of maths anxiety. During a two-week post-primary school placement in my first term of doctoral training in Educational and Child Psychology, I also witnessed these

experiences of TA. At the time, students were facing standardised testing in numeracy and literacy, specifically the Drumcondra Post-Primary Tests. As a consequence of this observation, I originally, and prematurely, assumed that addressing this issue would be most necessary in post-primary schools. I then discovered that TA had become a topical issue at the earlier stage of primary-level. This realisation stemmed from conversations with lecturers who gave examples of primary-aged children who experienced a decline in performance in standardised testing due to extreme anxiety, as well as from friends who are primary teachers and noted the prevalence of this issue. The awareness of TA in this age group was consolidated by my reading into the reported concerns of Irish primary-level teachers in relation to TA (Devine et al., 2020; O'Leary et al., 2019). Consequently, this piqued my interest in learning about what can be done to alleviate this issue, in other words, effective TA prevention and intervention.

1.4. Research Questions

1.4.1. Primary Research Question

- How effective is Weems' (2015) school-based TA intervention as a universal programme in reducing TA in a sample of Irish primary-level children?

1.4.2. Secondary Research Questions

- What effect does this intervention have on children's perceived self-efficacy of academic achievement?
- What effect does this intervention have on children's emotion regulation skills?

1.5. Thesis Structure

Following this introductory chapter, the thesis is presented in three chapters, a review paper, an empirical paper, and a critical review and impact statement.

Chapter Two: Review Paper

This chapter provides information on the origins of anxiety, childhood anxiety, and more specifically, TA. A broad theoretical history of TA is discussed, followed by an overview of the issue and prevalence of TA in a primary setting. This section emphasises the importance of effective TA intervention in the school setting at primary-level. Subsequently, a critical and systematic review of the evidence is detailed. The conclusion of this review highlights a gap in the literature and consequent rationale for the main study.

Chapter Three: Empirical Paper

Chapter three briefly reiterates the issue of TA and aims of the main study in light of the findings from the systematic review. Primary and secondary research questions, and their corresponding hypotheses, are then outlined before detailing the methodology of the study. The methodology section includes an overview, theoretical background, and evidence-base for the chosen intervention. This is followed by a detailed outline of the results from statistical analysis of the gathered data. Finally, a discussion of the interpretation and implications of these results are presented, including references to the study's strengths and limitations and possible explanations of the findings.

Chapter Four: Critical Review and Impact Statement

The final chapter of this thesis critically appraises the research and provides a personal reflection on the overall research process. This includes a critical discussion of the epistemological and ontological perspective adopted. Additional strengths and limitations of the study are discussed in further detail and in relation to the researcher's decision-making process. This chapter also presents a detailed overview of the implications of the findings in relation to TA intervention, educational policy and curriculum, practice, and future research, as well as an overview of the distinct contribution of this research.

Chapter Two: Review Paper

This chapter provides a detailed overview of the key issue of test anxiety (TA), followed by a systematic and critical account of literature related to school-based TA interventions at primary-level. Subsequently, this review devises a context and rationale for the main study.

2.1. Introduction to Anxiety

2.1.1. Definition

The American Psychological Association (2021) defines anxiety as an emotion characterised by worried thoughts, feelings of tension, and physiological changes, such as, increased blood pressure and rapid heartbeat. It is a complex state of psychological distress in the face of threatening stimuli which elicits emotional, behavioural, cognitive, and physiological reactions (Weis, 2020). Signs of anxiety may include ‘somatic’ symptoms, such as sweating, muscular tension, and dizziness, and ‘psychological’ symptoms, such as irritability, feelings of dread and threat, panic, and worry (Talley & O'Connor, 2010).

2.1.2. Origins of Anxiety

It has been stated that the main characteristic feature of anxiety is ‘worry’ (Baldwin & Leonard, 2013). Of course, certain levels of fear, worry, or anxiety are normal and have long been recognised as necessary for survival (Williams & Knight, 1994). Historically, these responses existed due to the evolutionary purpose of supporting safety in the face of threat (Grant et al., 1994) by stimulating alertness to prepare the body for fight or flight (Marques & Metcalf, 2013). At this point, the distinction between fear and anxiety must be addressed. Fear is considered to be a response to a real event, whereas anxiety is associated with the anticipation of an event, which may or may not occur (Levine & Munsch, 2016). To clarify this difference further, Chapman et al., (2011) present the example of a rollercoaster ride whereby you may feel anxious in anticipation of a steep slope and ultimately experience fear as the rollercoaster descends rapidly down the slope.

Therefore, it is acknowledged that fear and anxiety are common responses to dangerous or uncomfortable situations (McLaughlin & Holliday, 2013). However, disordered anxiety differs from normal feelings of anxiety which occur at times of stress, worry, or real danger (American Psychiatric Association [APA], 2015). In contrast, anxiety levels are considered to be ‘disordered’, or ‘problematic’, when there is no threat and they are no longer adaptive (Bagnell, 2011), or similarly, if the fear response is excessive and/or abnormal given

the situation (Robichaud et al., 2015). Furthermore, extreme anxiety is characterised by an impairment in daily life, including adaptive, personal, social, and/or occupational functioning (Edmunds & Mayhew, 2000).

2.1.3. Childhood Anxiety

Correspondingly, these facts regarding anxiety also apply to the average child's developmental trajectory of anxiety. Fears and worries are a typical part of development and their specificity differs throughout childhood and adolescence (Bagnell, 2011). For example, young infants commonly fear loud noises, with fears of strangers and separation arising at the end of their first year (Stallard, 2014). Preschool-aged children typically fear separation, monsters, animals, and storms (Bagnell, 2011). Older school-aged children generally begin to worry more about injury and death (Hockenberry & Wilson, 2014). At the point of adolescence, worries tend to deviate towards concerns over social comparisons and fears regarding academic tests, failure, and criticism (Warren & Sroufe, 2004). Despite the regularity of these experiences, when anxiety becomes so severe that it interferes with the child's functioning, including at home, at school, and/or other social settings, it is then considered an issue of significant anxiety (Cooper, 2012). Furthermore, if the fear is not congruent with their developmental level, this may also be indicative of interference in functioning (Essau & Ollendick, 2012). For example, a fear of separation in school-aged children is not appropriate and can hinder social functioning (Phares, 2003).

Significant anxiety can impact children of all ages and can manifest in a variety of forms (Rapee et al., 2008). The Diagnostic and Statistical Manual, Fifth Edition (DSM-5; APA, 2013) classifies several distinct anxiety conditions including specific phobias, selective mutism, generalised anxiety, agoraphobia, separation anxiety, panic disorder, and social anxiety or phobia. The present research is focused on the distinct category of TA. Although not recognised in the DSM-5, it has been argued that TA is most similar to, or falls under the umbrella of, social anxiety, and was previously mentioned under this category in the DSM-IV (APA, 1994). This comparison has been drawn due to the fundamental fear of negative social evaluation which is at the core of TA (Schneier & Heckelman, 1995). The concept of TA was originally considered for inclusion in the DSM-5, however, due to difficulty in defining it and an excessive proportion of the population which it captured, it was excluded from the manual (Bögels et al., 2010).

2.2. Overview of Test Anxiety

2.2.1. Definitions and Theories of Test Anxiety

As discussed, there is a multitude of specific forms of anxiety which can present in childhood (Rapee et al., 2008). TA is one type of anxiety which results from the pressure of testing situations, or circumstances where a person believes they are being evaluated (Cizek & Burg, 2006). It is described as an extreme fear of assessment or evaluation settings (Nata, 2007) which has a debilitating effect on performance (Larsen, 2017). TA is distinguished from generalised anxiety in this sense as research shows that TA is significantly correlated with test performance, whereas generalised anxiety is not (Sarason & Palola, 1960).

The Origins and Early Theories of Test Anxiety. As a concept, TA was originally theorised in 1952 by George Mandler and Seymour Sarason in order to explain the effects of anxiety on performance in testing situations (Hasija, 1993). Mandler and Sarason (1952) proposed two types of learned drives which occur in testing situations, one for task performance and one for anxiety. More specifically, they postulated that a ‘learned task drive’ transpires whereby the demand characteristics of the task stimulate task responses and reduce the anxiety drive through task completion (Zuckerman et al., 2015). The second drive, ‘learned anxiety drive’, can elicit either task-relevant responses, or task-irrelevant responses which interfere with performance (Fisher, 2015). These task-irrelevant responses are attributed to “feelings of inadequacy, helplessness, heightened somatic reaction, anticipation of punishment or loss of status or esteem” (Mandler & Sarason, 1952, p.166). This theory was later criticised for attributing TA to sole underlying causal entities, in the form of ‘drives’, which are unobservable and too obscure to target for reduction (Richardson, 2014).

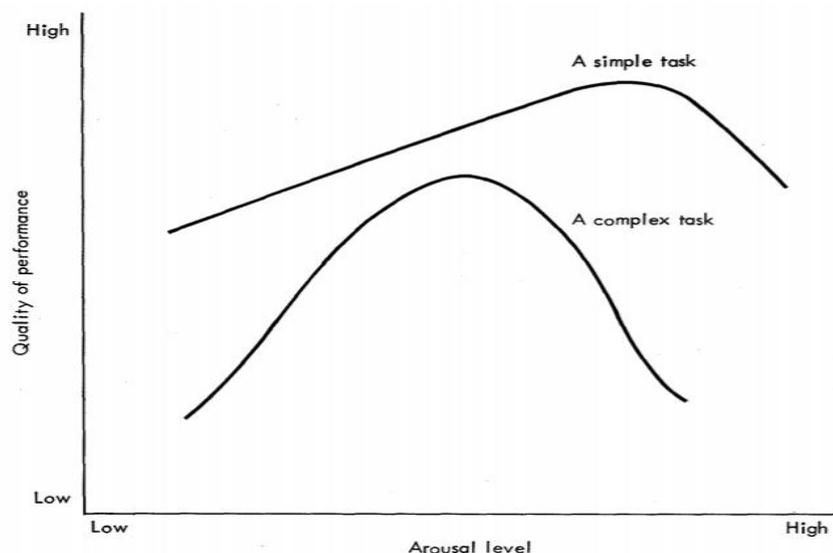
In the early 1960s, researchers then categorised TA as ‘facilitating’ or ‘debilitating’ based on the varying degree of severity (Alpert & Haber, 1960). This differentiation has been recognised as an important contribution to the literature on TA, both conceptually and methodologically (Zeidner, 2008). To elaborate on this distinction, nervousness is a normal reaction to test situations and can facilitate learning due to increased motivation to study (Larsen, 2017) and enhanced arousal for concentration in classroom and test settings (Rosenthal, 2013). For example, in a classroom setting, a certain degree of TA motivates students to focus, take notes, and attend to class activities, and in testing situations prompts careful reading, proofreading, and other positive test-taking strategies (Sullivan, 2009). Consequently, anxiety can enhance test performance if it is minimised and controlled to maintain facilitating effects (Payne, 2020). Beyond this point, is the onset of debilitating

anxiety, which is experienced by highly test-anxious individuals (Tryon, 1980). It is referred to as ‘debilitating’ as a result of the negative impact on performance whereby the individual’s fear of failure can hinder concentration during class activities, study, and/or testing situations (Sullivan, 2009). In more recent years, it has been argued that it is only at this point, when anxiety becomes so extreme that it interferes with performance, that it is classed as TA (Larsen, 2017).

Yerkes-Dodson (1908) law was also employed in early literature to illustrate this spectrum of positive to negative effects of TA. This law proposes an inverted U-shaped function of arousal whereby quality of performance is dependent on arousal levels which vary based on task complexity. Based on this law, a level of anxiety midway through the curve elicits optimal performance (Marques & Metcalf, 2013). Beyond this optimal level, towards the peak of arousal, anxiety becomes maladaptive due to distress and consequent impairment of functioning (Buchwald & Schwarzer, 2011). An illustration of Yerkes-Dodson law (Kahneman, 1973, p.34) is displayed in Figure 1.

Figure 1

Yerkes-Dodson (1908) Law as cited in Kahneman (1973, p.34)



Despite illustrating the association between arousal and performance, this theory’s shortcoming as a mere description has been discerned (Davies et al., 2013). It has been criticised for failing to provide any information explaining why this anxiety-performance relationship occurs (Zeidner, 2006).

Furthermore, there is now broad agreement that TA is conceptualised almost exclusively as debilitating (McDonald, 2001). Therefore, these theories of facilitating and debilitating

effects were not widely adopted, despite the important contribution of Alpert and Haber's (1960) theory.

The Emergence of Cognitive and Attentional Theories. As mentioned, Mandler and Sarason's (1952) pioneering theory of TA was criticised for focusing solely on mechanistic drives which disregarded the cognitive and attentional dynamics of TA (Richardson, 2014). In the late 1960s and 1970s, these cognitive and attentional variables came to the forefront of TA literature, in line with a cognitive revolution (Zeidner, 2004).

Most notable was the introduction of Liebert and Morris' (1967) supposition that TA is comprised of two key concepts, 'worry' and 'emotionality'. The 'worry' component refers to the cognitive activity of concern regarding performance, whereas 'emotionality' represents the physiological reaction to anxiety (Sapp, 2013). These two elements were shown to be correlated whilst remaining empirically distinct (Liebert & Morris, 1967). This distinction was highly influential in shifting TA theory towards a more cognitive orientation (Zeidner, 2008). This was followed by Spielberger et al. (1978) who refined TA into a measurable two-dimensional construct using the concepts of worry and emotionality. Spielberger (1966) had also previously provided an important distinction between two general anxiety concepts, namely 'state' and 'trait' anxiety. State anxiety refers to the immediate feelings of anxiety in a given moment, whereas trait anxiety refers to a stable personality characteristic of being anxiety-prone (Zeidner & Matthews, 2010). Based on this state-trait theory, TA was conceptualised as a situation-specific anxiety trait (Spielberger et al., 1976). It should be noted, however, that although anxiety is related to the evaluative situation and has the most debilitating effect in this state, TA can also extend beyond the testing situation to the weeks leading up to and following the test (Coomer & Latham-Mintus, 2019). Based on these concepts, it was posited that both TA-induced worry and emotionality impact academic achievement (Hodapp et al., 1995) whereby the TA construct of 'worry' is said to be comparable with trait anxiety, and 'emotionality' corresponds to state anxiety (Sapp, 2013).

Another important cognitive theory was proposed by Wine (1971) who presented evidence emphasising the importance of the direction of attention during testing. According to this attentional theory, test-anxious individuals divide their attention between task-relevant efforts and preoccupations with self-criticism, worry, and somatic concerns. As a result, performance suffers due to a lack of attentional units available for task-directed activities (Eysenck, 2015). This means that highly test-anxious individuals are more likely to be distracted from the cognitive task of a test compared to their low-anxious counterparts, who are more focused on the task (Wine, 1980). This theory was the first explicit interpretation of

TA based on cognitive attention (Allen, 2016).

In later years, Sarason (1984) expanded the two-dimensional TA frameworks previously outlined to propose a four-dimensional model of TA. The additional two factors included bodily symptoms and task-irrelevant thinking. It was proposed that the worry component of TA was comprised of bodily symptoms including bodily arousal and tension (Sapp, 2013), which, along with emotionality and task-irrelevant cognitions, devised a four-factor model. Sapp (1991) also found two concepts which correspond to the multiple dimensions proposed by these theories of TA. He conceptualised TA from the outlook of cognitive-behavioural theory whereby he viewed the cognitions and affects that stem from TA as behavioural responses (Sapp & Farrell, 1994). Therefore, from this perspective, it was suggested that worry corresponds to the cognitive component and emotionality represents the consequent behavioural responses (Sapp, 2010).

These theories of TA conceptualise an interference model whereby TA impairs performance by disrupting an individual's ability to recall learned information (Hembree, 1988). In other words, although a person may have sufficient knowledge and understanding of the content for a test, TA causes distraction due to worry and task-irrelevant cognitions and, therefore, their mind goes blank (Bin Kassim et al., 2008; Buchwald & Schwarzer, 2011).

An alternative to this interference model is the deficits model of TA (Naveh-Benjamin et al., 1987). Based on this model, a lack of preparation or ineffective study causes TA and performance failure (Buchwald & Schwarzer, 2011). In addition to poor study skills, this model also considers deficient test-taking skills as a contributing factor (Bin Kassim et al., 2008). In both cases, the individual is aware of previous poor performance and, therefore, TA is heightened (Ashcraft, 2019). Despite the established association between poor study habits and TA, this theory does not account for the onset of TA in high-achieving students with appropriate study skills (Tobias, 1985).

Subsequently, the 'dual deficit' or 'information processing' model aimed to bridge the gap between these interference and deficits models (Stroud, 2013). According to this theory, both task-irrelevant thoughts and skills-deficits can act as contributors to TA (Jones & Petrucci, 1995). This theory is supported by empirical findings which suggest that highly test-anxious individuals perform poorly due to cognitive interference, and/or poor study skills (Naveh-Benjamin, 1991).

Contemporary Theories. Beginning in the mid-2000s, several theories which offered additional and alternative explanations for the effects of TA emerged. Of significance was

Zeidner and Matthews's (2005) theory which emphasised a distinction between TA as an attribute of the individual and as a dynamic process. Based on the first element of this theory, dispositional TA may be understood as a contextualised personality trait (Zeidner & Matthews, 2010). Differences in personality may include the disposition to react with excessive worry, mental disorganisation, and physiological arousal in testing situations. Accordingly, individual attributional differences can account for variation in experiences of TA, and subsequent beliefs and behavioural responses (Cassady, 2010). Secondly, in relation to the process-oriented perspective, TA also depends on the reciprocal interaction of several distinct factors. This includes the evaluative context, the individual's vulnerability to anxiety, threat perceptions, appraisals, coping skills, and adaptive outcomes (Carducci et al., 2020). Drawing on these interactive factors, Zeidner and Matthews (2005) proposed six types of TA which were corroborated by previous research on TA theories including: a) study or testing skills deficits, b) anxiety blockage or retrieval failure, c) failure-acceptance whereby anxiety nurtures learned helplessness and reduced motivation, d) failure-avoidance which is the avoidance of testing situations, e) self-handicapping behaviours which involve preserving self-worth and ego through self-imposing barriers to success and f) maladaptive perfectionism which involves excessively high performance expectations.

Another theory, which expands on some of the discussed TA theories above, is Pekrun (2006) and Pekrun and Perry's (2014) Control Value Theory of Achievement Emotions. Although not a TA theory specifically, it can be considered here as TA falls under one of the proposed achievement emotions (Putwain & Synes, 2020). Extending on previously proposed theories of emotions pertaining to failure and success outcomes, Pekrun and Perry consider a broader range of achievement emotions. This variety of emotions are characterised into two forms: outcome emotions and activity emotions. They are determined based on individual appraisals of controllability over activities and outcomes which are subjectively important (Pekrun, 2017). Variations of perceived control and value appraisals are said to elicit different types of achievement emotions (Boekaerts & Pekrun, 2015). TA is related to outcome emotions as this category includes feelings of hopelessness and anxiety due to the person's belief that they cannot prevent failure in an important exam. Therefore, the extent of TA is determined by their estimation of risk, expectancy of failure, and/or subjective value of success (Baldwin & Leonard, 2013; Pekrun, 2016). Expectancy of failure is said to be reliant on self-efficacy, and the appraised value is dependent on personal achievement goals, therefore, subsequent TA can be determined by these factors (Pekrun, 2017). In examination of this control value theory, several studies have found that both perceived control and value

are significant predictors of outcome emotions, including anxiety (Pekrun et al., 2011; Turner & Schallert, 2001).

In extension of the theoretical literature, a biopsychosocial model has recently been proposed to include a ‘social’ component of TA (Flippo et al., 2018). Lowe et al. (2008) proposed that TA is influenced by social factors, whereby a person may fear failing a test due to potential social humiliation or negative responses from significant others, such as peers, parents, and/or teachers. Therefore, the TA factors included in this biopsychosocial model are categorised as ‘biological’ which is the physiological response, ‘psychological’ which refers to the emotional and cognitive factors, and the ‘social’ component as outlined above, as well as task-irrelevant behaviours (Von der Embse et al., 2014). This social dimension of TA is supported by empirical studies (Lowe, 2014; Lowe et al., 2011).

Finally, and most recently, Segool et al. (2014) built upon this biopsychosocial model using advanced statistical modelling to examine a cognitive-behavioural framework of TA. In this study, they identified the combination of cognitive perceptions, and prior academic experiences with demographic factors, social or educational context, and educational expectations as influencers of TA (von der Embse et al., 2018).

Theoretical Conclusions. In closing this section on the history of diverse TA theories, it is clear that TA is a complex phenomenon with a range of components and effects (Anderson & Sauser, 1995). Upon review of these theories, it is difficult to agree on one position due to the merits and limitations attributed to each. However, current trends agree that TA is a multidimensional construct and presents in a variety of forms (Stöber & Pekrun, 2004). Furthermore, there is broad agreement that impaired academic performance and subsequent self-efficacy are core associates of TA, which can act as both antecedents and consequences. In conclusion, it is evident from the breadth of theoretical literature that a multitude of contributing elements and effects of TA exist. These elements will be discussed in relation to empirical causes and correlational factors later in this chapter.

2.2.2. Manifestations of Test Anxiety

Similar to generalised anxiety, and as described in the theories outlined above, TA can manifest in a variety of forms including cognitive, behavioural, physiological, and emotional symptoms (Zeidner & Matthews, 2005). The various categorised symptoms of TA are outlined in Table 1 (Kennedy, 2010; Sawka-Miller, 2011).

Table 1*Symptoms of Test Anxiety* (Kennedy, 2010; Sawka-Miller, 2011)

Classification	Examples
Cognitive	Worry, negative self-talk, attention and concentration difficulties, and memory difficulties.
Behavioural	Task avoidance, withdrawal, freezing, and fidgeting.
Physiological	Increased heart rate, headaches, muscle tension, sweating, rapid breathing, shortness of breath, nausea, shaking, dry mouth, fainting, vomiting, diarrhoea, difficulty sleeping, and panic attacks.
Emotional	Feelings of hopelessness and inadequacy, anger, fear, low self-esteem, low confidence, low self-efficacy, and depression.

This table outlines an extensive list of potential symptoms of TA, meaning that not all test-anxious individuals will experience every symptom. Instead, they may experience a variety of combinations of symptoms. These symptoms can manifest before, during, and/or following a test (Coomer & Latham-Mintus, 2019).

TA can also present in the context of generalised anxiety and/or social anxiety disorders (Simos & Hofmann, 2013), with some arguing that TA is an element of general anxiety (Sapp et al., 1995; Sieber et al., 2013). Individuals who experience generalised anxiety worry about everyday matters (Holloway, 2019). Therefore, children with generalised anxiety understandably worry more about tests, compared to their non-anxious peers (Brown, 2021). This is supported by research which has shown that children with higher levels of TA also report higher levels of general worries (Beidel & Turner, 1988). Similarly, TA can stem from social anxiety due to the nature of worries associated with the disorder (Heimberg & Barlow, 1995). Specifically, social anxiety involves an extreme fear of being negatively judged in social or performance situations, including school tests (Dugas & Robichaud, 2007). These tests can be particularly anxiety-inducing for individuals who desire to make positive impressions but fear that their results will not align with their aspirations (Leary & Kowalski, 1997). Although research demonstrates that TA can overlap with alternative anxiety disorders, the similarities are not sufficient to make these concepts synonymous (Sieber et al., 2013). TA has been shown to be a distinct construct (Flippo et al., 2018) and

can present in the absence of other anxiety disorders (Antony & Norton, 2015). Some individuals who are typically relaxed in their everyday lives can experience extreme anxiety in testing situations (Zeidner & Matthews, 2010). This can occur for several reasons, such as, previous failure, lack of preparation, or poor study skills (Naveh-Benjamin, 1991; Putwain, 2008; Silvestri & Silvestri, 2013). It has also been argued that TA differs from the concept of generalised anxiety in that it defines a specific context in which anxiety occurs (Putwain, 2008), whereas generalised anxiety is unrelated to specific matters (Rickels & Rynn, 2001). Generally, it can be difficult to distinguish between anxiety-based issues as individuals often experience symptoms related to various anxiety disorders, rather than exclusively aligning with one category (Antony & Norton, 2015). In conclusion, TA can manifest as an element of different anxiety disorders in some, whilst presenting as a distinct problem in others. Regardless of the context in which it occurs, TA requires targeted intervention to prevent adverse effects; such consequences are detailed later in this chapter.

2.2.3. Test Anxiety: Contributing and Correlational Factors

Literature surrounding TA and its associations have identified a variety of contributing and casual factors which can interact to increase the likelihood of TA.

Firstly, some individuals have a stable disposition to experience anxious states, often observed in a neurotic personality type (McDonald, 2001) and reflective of ‘trait’ anxiety (Spielberger, 1966). As previously discussed, TA can correlate with generalised anxiety, whilst notably remaining a distinct construct, suggesting that those with general anxiety are more likely to experience TA (Flippo et al., 2018). For some individuals, this susceptibility to anxiety may be attributed to a biological or genetic predisposition (Huberty, 2009). In other words, if there is a family history of anxiety conditions and/or neurotic personality type, this may be transmitted to the child through their parent (Beidel & Alfano, 2011). Interestingly, TA has also been shown to be consistently higher in females compared to their male counterparts, indicating a gender vulnerability to TA (Putwain & Daly, 2014; Putwain, 2007).

Literature surrounding the causal factors of TA have also identified appraisals concerning threat of failure which elicit worry (Pekrun, 2011). As suggested by the deficit model (Naveh-Benjamin et al., 1987), this worry can be heightened by a lack of test preparation (Silvestri & Silvestri, 2013). An increased fear of failure can also stem from previous experiences of failure (Putwain, 2008). It is important to note that ‘failure’ can be objective, whereby failure is defined as not passing the test, or subjective, whereby failure may be viewed as getting a grade of a B or C based on individual expectations (Dryden,

2011; Van Blerkom, 2013). These high expectations or personal standards, may be referred to as 'perfectionism' (Schuler, 2000). Unsurprisingly, it has been found that maladaptive perfectionism, which is characterised by high-performance expectations and extreme self-blame when failing to meet desired standards, is a significant predictor of TA (Eum & Rice, 2011). Moreover, research on the link between motivation and TA has demonstrated that socially-prescribed perfectionism is positively correlated with TA, whereas self-oriented perfectionism is not (Stoeber et al., 2009). This suggests that motivation to impress others can heighten TA due to a fear of failure and/or disappointment (Olatoye, 2009). More generally, subjective test importance has been shown to influence test motivation, and subsequently increase TA (Cheng et al., 2014). In addition to test importance and motivation, experiences of subjective and/or objective failure can result in heightened TA due to diminished confidence and self-efficacy (Marcz, 2017). Self-efficacy is one's judgement of their own capabilities needed to attain certain types of performances (Bandura, 1986). Therefore, low confidence and academic self-efficacy increase the tendency towards the onset of TA due to threat of failure (Banks, 2012; Khan & Madden, 2018). An individual's self-efficacy can differ in each school subject, in other words, one may judge themselves as highly capable in one subject but not in another (Putwain, 2008). Correspondingly, individuals with low self-efficacy in certain academic subjects are more likely to experience TA related to tests in those subjects (Banks, 2012). Although it is not the only subject to induce TA, mathematics is generally judged to be the most anxiety-provoking subject in schools (Dowker, 2019). This has been attributed to parents' and teachers' maths anxieties unintentionally influencing children's perceptions of maths (Durwin & Reese-Weber, 2019; Smith, 2021), the high-stakes nature of standardised maths tests (Sousa, 2016), and the stereotype that maths is difficult to do well in (Boaler & Dweck, 2015). As such, the literature suggests that one's personal competence across academic subjects should also be considered when examining contributing and maintaining factors of an individual's TA (Vargios, 2007). In addition to subject-specific TA, the nature of evaluation settings has also recently been considered to elicit variation in TA levels. For example, heightened TA may occur in those who experience TA due to fear of social humiliation when facing oral tests as they involve a greater social component, when compared with written tests (Sparfeldt et al., 2013). This is supported by research which found that oral TA is related to social anxiety, whereas TA for written exams is not (Laurin-Barantke et al., 2016). On the other hand, children with literacy and/or writing difficulties tend to perform better on oral assessments rather than written (Riddick, 2009),

and therefore, these modes of assessments may be less anxiety-inducing for this population, although this has yet to be researched.

As reflected in the biopsychosocial model of TA, environmental and cultural factors can also contribute to TA. Firstly, parental pressure and attitude have been directly linked to TA (Zeidner, 2014). This occurs whereby children internalise TA dispositions due to high parental demands and expectations (Kaya, 2004). Unsurprisingly, this parental pressure to achieve high standards has also been associated with a perfectionist personality trait in parents (Casbarro, 2005). Furthermore, family stressors, such as, parents fighting, financial stressors, and sibling rivalry, can also contribute to TA (Yew Chye, 2008). Additionally, teacher stress related to pupils' test results can filter down to children who subsequently feel pressure to maintain a certain standard of achievement (Carr, 2015). Competition among peers, which is more prevalent amongst females, can also contribute to this stress (Bala & Shaafiu, 2016). It has been reported that these pressures are heightened in cultures where there is a higher emphasis placed on optimal achievement, such as in Asia (Chen & Kaspar, 2004). This is clear from a cross-cultural study which reported that Asian and American cultures have an inflated prevalence of TA relative to their European counterparts (Sharma & Sud, 1990).

Finally, individuals from a range of minority groups are more likely to confront TA. This includes children from racial and ethnic minority backgrounds (Baron et al., 2003). It has been suggested that these individuals experience TA stemming from a fear that their test performance will conform to the stereotype that they may be less intellectually capable (Frisby, 2013). Similarly, research has shown that individuals from low socio-economic backgrounds report higher levels of TA (Putwain, 2007). Finally, individuals with intellectual and learning disabilities experience greater rates of debilitating TA than those without disabilities (Datta, 2014; Whitaker Sena et al., 2007). These learning issues cause stress and difficulty in relation to academics and, therefore, give rise to a sense of threat in test situations (Dendy et al., 2003; Pierangelo & Giuliani, 2008).

2.3. The Issue of Test Anxiety

Global anxiety issues have become increasingly prevalent in children and adolescents in recent years (Bushnell et al., 2020), including a steep rise discovered in an Irish study (Dooley et al., 2019). One of the most obvious examples of childhood anxiety is TA and anxiety related to doing well in school (Casbarro, 2005). This type of anxiety is a universal

phenomenon experienced by individuals across the globe in evaluation settings or events (Bin Kassim et al., 2008). Sterian and Mocanu (2013) recognised TA as one of the most common types of anxiety and emphasised the importance of appropriate interventions.

As previously mentioned, the onset of fears and anxiety over evaluation and achievement were traditionally associated with adolescence (Warren & Sroufe, 2004). However, concern over the increasing prevalence of anxiety related to testing in primary-level children has emerged more recently.

2.3.1. Increasing Prevalence of Test Anxiety at Primary-level

In recent years, TA has been recognised as a growing issue at primary-level due to an increased focus on standardised testing fostering a culture of anxiety in schools (Lobman, 2014). This onset of TA can begin as early as seven-years-old (Von Der Embse et al., 2013) with increased levels observed in grades four to five (Lowe, 2019), equivalent to fourth to fifth classes in Ireland. In an Irish context, one major survey reported that 50% of primary-level teachers agree that standardised testing causes performance anxiety for children (Devine et al., 2020). It is worth noting, however, that whether it was a healthy facilitating anxiety or extreme debilitating anxiety was not disclosed. According to another Irish survey, teachers reported pupils experiencing extreme anxiety prior to standardised tests (O'Leary et al., 2019), with a higher number of three out of four teachers agreeing with this. Furthermore, in a study of self-reported TA in primary-level children, it has been found that this type of high-stakes testing elicits a heightened prevalence for moderate TA compared to classroom testing, which is more likely to induce low-level TA (Segool et al., 2013). This may be attributed to the pressure placed on schools and teachers to maintain success in pupils' results (Glazzard & Bancroft, 2018). These results are often seen as a reflection of their quality of teaching (Carr, 2015). Furthermore, it is mandatory in Ireland for standardised test results to be reported to parents (Department of Education and Skills [DES], 2006, 2011); this accountability can be an additional source of stress for teachers (Mathison & Freeman, 2006). Consequently, expectations from schools, as well as parents, can place pressure and stress on pupils to succeed (Allen & Kern, 2017). Interestingly, however, Segool et al.'s (2013) study found that a similar number of participants reported high TA across high-stakes standardised test and classroom test conditions. This would suggest that classroom tests may be just as anxiety-provoking for highly test-anxious individuals.

Based on these striking reports, it seems logical that TA interventions should be targeted at primary-level to ensure that children learn appropriate coping mechanisms at a

young age. In turn, this would also prepare students for the demands of tests and examinations at post-primary and third-level education to prevent increasing TA. This is particularly important given the increasing frequency of testing which a child faces as they progress through the educational system, accompanied by greater pressure and expectations (McDonald, 2001). This continued and increasing pressure is evident from reports of extreme stress and mental and physical health issues, as a result of exam pressure, by Irish Leaving Certificate students (O'Riordan, 2019). This is likely due to the high stakes of these examinations in determining students' futures, such as entry to college and subsequent careers (Zamudio et al., 2011). Furthermore, pressure to achieve in examination settings does not appear to ease at third level, as evidenced by reports of TA by university students (Chapell et al., 2005). This highlights the need for early proactive intervention to prevent these issues in later life.

2.3.2. Consequences of Test Anxiety

TA has been cited as a pivotal factor in determining an array of negative outcomes, such as, academic underachievement, impaired cognitive performance, psychological distress, and health issues (Zeidner, 2007). There is vast evidence and agreement amongst researchers that TA negatively impacts performance (Cassady & Johnson, 2002; Chapell et al., 2005; Rana & Mahmood, 2010), including at primary-level (Crişan & Copaci, 2015). As a result, TA has been condemned as one of the most disruptive contributing factors to academic underachievement (Bin Kassim et al., 2008). This means that test-anxious individuals often do not accomplish results reflective of their true potential (Siegle, 2018). Understandably, this can cause disappointment, embarrassment, and/or frustration (Van Blerkom, 2013). Furthermore, chronic underachievement, as a result of TA, can create a vicious cycle whereby TA further increases due to these negative experiences, feelings, and fear of re-current failure (Cizek & Burg, 2006). This is due to the mutual association of academic achievement and self-efficacy (Olivier et al., 2019), whereby an individual's belief in their own capabilities is diminished as a result of objective or subjective failure (Kratochwill et al., 2003). To reduce this fear of failure, some individuals may spend excessive time studying, whilst others may resort to cheating (McMillan, 2017). Alternatively, a chronic lack of reward, corresponding to the amount of effort put in, can lead to diminished motivation and effort (Cizek & Burg, 2006). In extreme cases, this can lead to test avoidance, a decrease in school attendance (Warner et al., 2018), and an increased likelihood of school drop-out (Vye et al., 2007).

The negative consequences of TA include not only hindered academic performance, but

also the distress of the young person (Spence, 1994). This distress can impact both physical and emotional wellbeing (Rothman, 2004; Schaefer et al., 2007). As mentioned, TA and impaired academic performance can negatively impact an individual's self-efficacy (Olivier et al., 2019). Consequently, feelings of self-doubt can lead to low confidence, self-esteem, and self-worth, and feelings of hopelessness (Damer & Melendres, 2011; Khan & Madden, 2018; Sari et al., 2018). More worrying is the significant and positive association with depression (Akinsola & Nwajei, 2013) and suicidal ideation (Lee et al., 2006). If not treated, TA can also lead to adverse health effects in the form of panic attacks (Kennedy, 2010) and physical health problems, such as, headaches, stomach and bowel issues, disordered eating, recurrent infections, and sleep difficulties (Edlin & Golanty, 2009). Sleep may also be impacted by the aforementioned excessive studying of people with TA, and can subsequently affect performance due to tiredness (Silvestri & Silvestri, 2013).

This extensive list of negative and serious consequences of TA highlights the importance of TA intervention and prevention. This is paramount given that the effects of TA can extend beyond the testing situation, including the preceding stress and worry, and feelings of anxiety experienced following the test (Coomer & Latham-Mintus, 2019).

2.4. Focus and Purpose of Review in an Educational Psychology Practice and Policy Context

Since the introduction of Circular 0138/2006 (DES, 2006), followed by Circular 0056/2011 (DES, 2011), standardised assessment of reading and mathematics in Irish primary schools has become compulsory in second, fourth, and sixth classes, with some schools administering them in additional class levels (National Council for Curriculum and Assessment [NCCA], n.d.). Furthermore, as outlined in the more recent Circular 0018/2022 (DES, 2022a), Ireland is taking part in an international assessment of mathematics and science skills in fourth class pupils in the year 2023. This was preceded by a field trial in 40 Irish primary schools between March to April 2022. The aim of standardised testing is to monitor pupils' progress to inform teaching and learning and to identify those who require additional support (DES, 2006). Information from these tests are deemed necessary given the vital role of literacy and numeracy in allowing pupils to access all areas of the curriculum (NCCA, n.d.). Similarly, all forms of pupil assessment, including class tests, have been described as an essential component in the evaluation of learning (Buchwald & Schwarzer, 2011). However, as TA can lead to impaired performance and consequent underachievement (Rana & Mahmood, 2010), it has the potential to jeopardise assessment validity (Zeidner,

2007). Consequently, these assessment results may lead to an inaccurate depiction of a child's abilities and progress (Siegle, 2018), therefore, rendering them redundant in meeting their purpose.

Given the pressure of these assessments, it is understandable that there is growing concern over TA in Irish children. It is vital to provide support in addressing this issue in the educational context, as this is where assessment situations and subsequent TA arise (Yeo et al., 2016). This provision of support is particularly important in adherence to the 'Wellbeing Policy Statement and Framework for Practice', (DES, 2018). This policy states that it is the role of the education system to teach children the knowledge and skills to prepare for challenges which may influence their wellbeing. This emphasis on wellbeing in education is further emphasised in a recent publication by the DES (2022b) which provides guidance for promoting wellbeing in schools. This means that teachers may have a pivotal role in accommodating children experiencing TA (Yew Chye, 2008). Importantly, Educational Psychologists (EPs) also have a role in this regard as they work consultatively to empower teachers to intervene effectively in support of all pupils' needs (National Educational Psychological Service [NEPS], n.d.). Therefore, it is imperative that EPs can advise and support teachers to minimise TA in schools based on best-evidence practice in their role as scientist practitioners (Birch et al., 2015). Furthermore, EPs have a role in direct intervention with pupils, as well as intervention training with teachers (Fallon et al., 2010). Subsequently, knowledge of effective intervention is vital to reduce the prevalence of TA and improve children's academic performance. In turn, this would allow teachers and EPs to gain a more accurate picture of pupils' learning outcomes and capabilities.

2.5. Existing Intervention Research

TA is an issue which has captivated the interest of researchers since the 20th century (Bögels et al., 2010). This includes a large body of research on intervention approaches. Unfortunately, there is limited evidence to inform EPs of effective TA intervention at primary-level. Previous reviews of TA intervention programmes, including a meta-analysis in 2003 (Ergene, 2003) and a systematic review (Von Der Embse et al., 2013) for the years 2000-2010, have concluded that behavioural and/or cognitive strategies are most effective in alleviating TA, particularly when combined with skills approaches (Ergene, 2003), as well as some promising evidence for biofeedback (Von Der Embse et al., 2013). This combination of intervention strategies, such as psychoeducation and relaxation training along with study skills teaching, has been recommended to address the complex dimensions of TA (Gregor,

2005). Despite strong evidence for these approaches, these systematic and meta-analytic reviews have recognised an extreme lack of research at primary-level. That being said, a more recent systematic review (Soares & Woods, 2020) identified several new primary-level studies. The authors noted, however, that the overall number of included studies was small and less than half of these included primary-level children. Furthermore, it was reported that there was a lack of replication of any of the included intervention studies to support generalisability across populations.

Due to this minimal research and given the growing concern over TA at primary-level, it is vital to systematically review and examine the recent literature of intervention studies for this age group.

2.6. Systematic Review of the Literature

2.6.1. Review Question

This systematic review examines the effectiveness of TA interventions at primary-level. The purpose is to inform EPs of best-evidence practice for reducing TA, as well as identifying outstanding issues in the literature which require further research. As previously mentioned, it is vital to support children experiencing TA within a school context, therefore, the focus of this review is school-based TA interventions. Data was collected by systematically searching and reviewing research studies which focus on decreasing TA levels at primary-level using targeted school-based interventions. This was inclusive of those delivered in individual, group, or universal settings. Relevant studies and their results were then examined to evaluate the evidence for school-based interventions.

2.6.2. Search Strategy

A systematic search of relevant journal articles was carried out in June 2020 using five online databases; Academic Search Complete, PsychINFO, PsychARTICLES, ERIC and British Education Index. A final search was also carried out in August 2021 to establish any additional publications in the interim. As part of these searches, a filter was applied to ensure that journals were peer reviewed and, therefore, a higher standard of quality (Thyer, 2008). In addition, searches were filtered for articles published between 2011-2020 and 2011-2021, respectively. This timeline was chosen prior to the initial search in June 2020 as the most recently published systematic review of TA intervention at this point had been conducted for the years 2000-2010 (Von Der Embse et al., 2013). The search terms outlined in Table 2 were used to systematically search the literature for relevant studies.

Table 2*Search Terms for Online Databases*

Outcomes		Education Level		Prevention
Test anxi* OR	AND	Primary school OR	AND	Interven* OR
Exam anxi* OR		Elementary school OR		Prevent* OR
Test worry* OR		Primary education OR		Program* OR
Exam worry* OR		Elementary education		Reduc*
Test stress* OR				
Exam stress* OR				
Test pressur* OR				
Exam pressur* OR				
Test strain* OR				
Exam strain* OR				

Note: * was used to ensure that all variations of a word are included. For example, 'prevent*' is inclusive of prevent, preventative, and prevention.

Searches were then refined by removing duplicates and screening, for set inclusionary and exclusionary criteria, of the abstract and title sections, and finally a review of the full articles. The inclusionary and exclusionary criteria are outlined in Table 3. Figure 2 illustrates the process of screening in the final systematic search to identify appropriate studies for answering the review question based on this criteria. Studies which were excluded based on these criteria are listed in Appendix A. Seven studies were identified from this search with an eighth study later discovered from additional reading. The final list of included studies is outlined in Appendix B with a summary of studies in Appendix C.

Table 3*Inclusion and Exclusion Criteria*

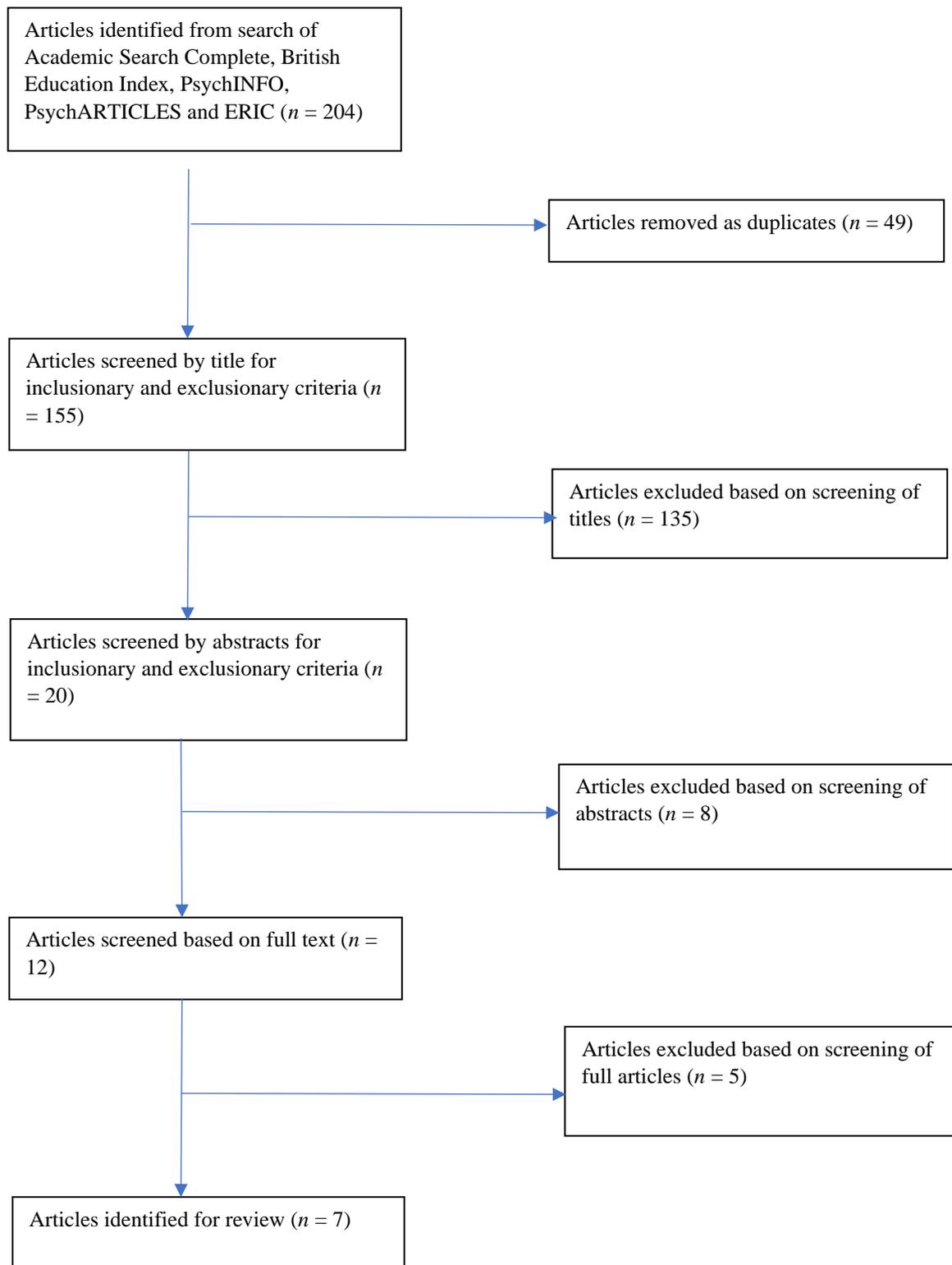
	Inclusion	Exclusion	Rationale
1. Type of anxiety	Test anxiety (TA)	General anxiety or other anxiety presentations	Focus of interest is on TA specifically
2. Intervention	School-based interventions	Interventions which are not school-based	As TA occurs in school settings, school is the most appropriate place for intervention
3. Years of publication	2011-2021	2010 and prior	A systematic review of TA interventions between 2000-2010 has previously been conducted (most recent review at the time of the original search). Circular 0056 (as previously discussed) was also published in the year 2011, marking an increased focus on standardised testing in Ireland
4. Type of publication	Peer-reviewed journal articles	Articles which have not been peer-reviewed	Peer reviews of studies means that they have been thoroughly criticised and are, therefore, likely to be of a higher standard
5. Population	Primary/elementary school-aged children	Individuals not attending primary or elementary level education	The focus of interest is on primary-level pupils
6. Language	Published in English	Any study not published in English	Practical for the understanding of the reviewer
7. Research design	Must include quantitative methods of data	Studies which do not include quantitative methods of data collection	This review examines the effect of interventions

collection and a
control group

Studies which do not
include a control group

on TA based on
empirical data

A control group must be
used for comparison to
determine how much of
the effects are due to the
intervention

Figure 2*Flow Chart of the Final Search (2021) Screening Process for Article Selection*

2.6.3. Weighting of Studies Framework

Gough's (2007) Weight of Evidence Framework was employed to thoroughly evaluate the evidence for effective TA intervention at primary-level. This framework is comprised of four sets of judgements for appraising evidence; Weight of Evidence A (WoE A) which is a generic judgement of the evidence, Weight of Evidence B (WoE B) which involves making a judgement on the methodology specific to the review question, Weight of Evidence C (WoE C) which judges the relevance of the focus of the study in answering the review question, and finally, Weight of Evidence D (WoE D) which averages WoE A, WoE B, and WoE C scores to provide an overall judgement.

Kratochwill's (2003) 'Task Force Coding Protocol' for group-based designs was utilised to evaluate WoE A as all included studies employed a group-based design. The protocol was amended based on the specific review question (See Appendix D for excluded sections). An example of the coding protocol for one study (Carsley et al., 2015) is illustrated in Appendix E. WoE B was then evaluated, followed by WoE C, and finally resulting in WoE D scores. Appendix D describes the process of scoring for WoE A, WoE B, and WoE C, including scores for each study. WoE scores are displayed in Table 4 below.

Table 4

Overall WoE Scores

Study	WoE A Score	WoE B Score	WoE C Score	WoE D Score (Average of WoE A, B, C Scores)
Carsley & Heath, (2019)	1.5 (medium)	3 (high)	3 (high)	2.5 (high)
Carsley et al. (2015)	1.25 (low)	3 (high)	2 (medium)	2.1 (medium)
Kurth et al. (2020)	1 (low)	0 (no evidence)	1 (low)	0.7 (low)
Mavilidi et al. (2014)	1 (low)	0 (no evidence)	1 (low)	0.7 (low)
Pourtaleb et al. (2018)	1.25 (low)	2 (medium)	2 (medium)	1.75 (medium)
Thompson et al. (2016)	1.25 (low)	0 (no evidence)	1 (low)	0.75 (low)

Weems et al. (2014)	1.75 (medium)	2 (medium)	2 (medium)	1.9 (medium)
Yeo et al. (2016)	2 (medium)	1 (low)	2 (medium)	1.7 (medium)

Note: Low = < 1.4, Medium = 1.5-2.4, High = > 2.5

2.6.4. Participants

The total number of participants for all studies in this review is 1815; this ranges from samples of 52 (Carsley et al., 2015) to 791 (Thompson et al., 2016). Low attrition was evidenced for five studies (Carsley & Heath, 2019; Carsley et al., 2015; Kurth et al., 2020; Mavilidi et al., 2014; Pourtaleb et al., 2018). Weems et al. (2014) experienced a 19% attrition rate from pre-intervention to post-intervention. Thompson et al. (2016) only reported the number of participants included in the final data analysis and, therefore, attrition rates were not disclosed. Reports of mean age of participants range from 8.7 (Kurth et al., 2020) to 11.59 (Mavilidi et al., 2014). Weems et al. (2014) did not disclose the mean age of participants but noted an age range of 8-17. Pourtaleb et al. (2018) reported that participants were sixth-graders. The gender balance varies among the included studies; five studies reported a balanced gender ratio with between 47-58% females (Carsley & Heath, 2019; Carsley et al., 2015; Kurth et al., 2020; Mavilidi et al., 2014; Thompson et al., 2016). Yeo et al. (2016) and Weems et al. (2014) reported a slightly less balanced ratio with 39% and 62% female samples, respectively. Pourtaleb et al. (2018) focused exclusively on a female sample. Two of these studies were conducted in Canada (Carsley & Heath, 2019; Carsley et al., 2015), with further studies carried out in Germany (Kurth et al., 2020), Greece (Mavilidi et al., 2014), Iran (Pourtaleb et al., 2018), the United States of America (USA; Thompson et al., 2016; Weems et al., 2014) and Singapore (Yeo et al., 2016). As the review question was focused on school-based interventions at primary-level, all studies included participant samples of children attending elementary/primary-level education. One study included both primary and post-primary pupils in their sample (Weems et al., 2014).

All included studies recruited participants by informing schools of the experiment and requesting voluntary participation of children with adult consent. The number of schools and classes which participant samples were taken from was considered when scoring WoE C to evaluate the generalisability of results. Six studies selected participants from various schools (Carsley & Heath, 2019; Mavilidi et al., 2014; Pourtaleb et al., 2018; Thompson et al., 2016; Weems et al., 2014; Yeo et al., 2016). Two studies selected participants from only one school each, however, participants were selected from various classes (Carsley et al., 2015; Kurth et

al., 2020). The type of school was also taken into consideration whereby four studies which were carried out in public school settings were given higher weighting (Carsley & Heath, 2019; Thompson et al., 2016; Weems et al., 2014; Yeo et al., 2016), as they are noted to be more likely to be in need of intervention in comparison to private schools (Kratochwill, 2003). One study carried out in a private school setting was allocated a lower score (Carsley et al., 2015). The type of school was unknown for three studies; therefore, they also received a lower score (Kurth et al., 2020; Mavilidi et al., 2014; Pourtaleb et al., 2018).

2.6.5. Interventions

All included studies involved the implementation of school-based TA interventions. Two of these studies incorporated mindfulness into an art activity which involved colouring a mandala for 15 minutes prior to a class test (Carsley & Heath, 2019; Carsley et al., 2015).

Kurth et al. (2020) also focused on a mindfulness-based intervention in the form of a short mindful breathing intervention. This involved completing a breathing exercise whereby participants were instructed to count their breath to 20 along with the researcher who counted aloud, prior to a test scenario.

Another study examined the effectiveness of a 'looking ahead' strategy in reducing TA (Mavilidi et al., 2014); this involved instructing participants that they had one minute to look through ten problems on a math test before allowing three minutes for solving each problem.

Pourtaleb et al. (2018) implemented an 'Integrated Training Programme'. The training package was comprised of a variety of different treatment methods, conducted during 14 40-minute sessions. This included 'behavioural intervention' in the form of progressive muscular relaxation and imaginative systematic desensitisation. This was followed by 'cognitive intervention' which involved recognising negative thoughts and cognitive restructuring. Finally, 'educational training' was provided, such as, study plans and methods, improving concentration for studying, and exam preparation.

Thompson et al. (2016) evaluated the effectiveness of a moderate to vigorous physical activity intervention delivered by school Physical Education (PE) teachers. Prior to sitting standardised tests for mathematics and reading, participants engaged in a highly active PE class which was based on a standardised protocol aligned with California's state physical education model content standards for fifth-grade pupils. Lessons lasted an average of 34

minutes, with an average of 15 minutes of moderate to vigorous physical activity. Lessons were comprised of a four-square grid, survivor tag, a water break, and fitness circuits, including 30 seconds of activity and 40 seconds rest, and switch intervals of running, skipping rope, and mountain climbers.

Weems et al. (2014) examined a group-based Cognitive-Behavioural Therapy (CBT) intervention which was delivered in five to six weekly sessions. The programme focused on psychoeducation followed by primarily behavioural strategies, including relaxation training and exposure techniques, with few cognitive strategies, such as, self-efficacy and self-evaluation training, as well as education on study and test-taking skills. The intervention was delivered as part of schools' counselling curriculums to fit with their ecology.

Finally, Yeo et al. (2016) also examined the effectiveness of a CBT intervention, however, they focused on a classroom-based approach. The programme incorporated behavioural strategy-focused intervention with cognitive modifications for targeting anxiety. This encompassed relaxation training, skills training, psychoeducation, self-instruction, and exposure to anxiety-provoking conditions. This intervention was built into curriculum time as a whole-class intervention and was carried out in four weekly 30-minute sessions in the month approaching the examination period.

2.6.6. Measures

Five of the studies used appropriate and reliable measures to determine TA levels. Two studies used Nilsson et al.'s (2012) children's version of the State Trait-Anxiety Inventory-State form to measure TA (Carsley & Heath, 2019; Carsley et al., 2015). This is a self-report measure of a child's current state of anxiety. Carsley and Heath (2019) reported reliability of $\alpha = .86$ for pre-intervention and $\alpha = .90$ for post-intervention, while Carsley et al. (2015) reported reliability of $\alpha = .82-.87$ for their sample. Pourtaleb et al. (2018) used the Spielberger (2010) Test Anxiety Inventory to measure TA levels; this is a self-report questionnaire, with reliability of $\alpha = .82$ for their study's sample. The Children's Test Anxiety Scale (Wren & Benson, 2004), which is also a self-report measure of TA in children, was used by Yeo et al. (2016); the reliability of this measure was reported as $\alpha = .94$ for this study. Weems et al. (2014) utilised a shortened version of the Test Anxiety Scale for Children (Sarason et al., 1958) with reliability of $\alpha = .82-.85$. Although reliable measures were used in each of these studies, it should be noted that TA was determined using only one measure and one source. This is reflected in their low WoE A score for measurements as multiple

measures or reporters of a primary outcome are recommended to ensure accuracy in detecting intervention effects (Mazurek Melnyk & Morrison-Beedy, 2018). Only two studies used multiple measures; firstly, Thompson et al. (2016) administered a self-report TA measure using the Children's Test Anxiety Scale, an adapted teacher version of the scale, and standardised maths and reading tests, which can be considered as an additional measure given the definition of TA as an extreme form of anxiety which negatively impacts performance. Mavilidi et al. (2014) also employed multiple measures in the form of a Greek translation of the Cognitive Anxiety Test scale (Cassady & Johnson, 2002) and maths tests to examine performance outcomes. Unfortunately, both studies employing multiple measures received a score of 'no evidence' for WoE A under the measurements category due to a lack of reported reliability of measures for the specific participant sample. Similarly, Kurth et al.'s (2020) study, which used a Smartband to measure physiological stress reactions, received a score of 'no evidence' under the same grounds.

2.6.7. Study design

All studies utilised quantitative methods to test for effects. Five of these studies compared TA levels within-subjects from pre-intervention to post-intervention, as well as differences between-groups (Carsley & Heath, 2019; Carsley et al., 2015; Pourtaieb et al., 2018; Weems et al., 2014; Yeo et al., 2016). Three studies did not compare TA levels at pre-intervention and post-intervention which is reflected in their lower scores for WoE B (Kurth et al., 2020; Mavilidi et al., 2014; Thompson et al., 2016). Instead, they conducted between-groups comparison of TA levels only. Although Thompson et al. (2016) compared test performance scores from pre-intervention to post-intervention, this was not their primary outcome. Furthermore, changes in performance over time cannot be attributed to TA when the same comparison has not been made for TA; this is important given that performance may be impacted by alternative factors, such as, health problems and chronic illness (Fertman & Grim, 2010), and school absences (Gottfried, 2013). Follow-up assessment at a later stage was carried out by Yeo et al. (2016) and Weems et al. (2014), however, the remaining six included studies did not complete follow-up measures resulting in no evidence scores for 'Follow-up Assessment' in WoE A. Although, Weems et al. (2014) conducted three follow-up measures, the study received a low score due to high attrition rates. All studies utilised a control group for comparison of TA outcomes. All studies, with the exception of one (Yeo et al., 2016), randomly assigned participants to either control or intervention conditions, showing reduced likelihood of differences between groups having influenced outcomes

(Balakrishnan, 2014). Yeo et al. (2016) reported that the school assigned participants based on convenience. The type of control group used for comparison also contributed to the studies' WoE A and B scores. Higher weighting was given to studies with 'active' control groups which involved an alternative intervention. Two studies exercised the use of active control groups in the form of non-mindfulness free-drawing conditions (Carsley & Heath, 2019; Carsley et al., 2015). Six studies used non-active (no intervention) control groups (Kurth et al., 2020; Mavilidi et al., 2014; Pourtaleb et al., 2018; Thompson et al., 2016; Weems et al., 2014; Yeo et al., 2016) resulting in lower scores.

2.6.8. Summary of Intervention Outcomes

From 2014-2020 there have been several studies which have examined various interventions and strategies for reducing TA; these studies have presented a variety of findings. Specific effect sizes have not been re-calculated as part of this review as studies differed in terms of sample sizes, measures, control groups, and designs. However, a synthesis of the findings is discussed in relation to the significance of effects and effect sizes reported individually in each study. Cohen's d effect sizes can be interpreted as .20 = small, .50 = medium, .80 = large, and eta squared (η^2) or partial eta squared (η_p^2) are categorised as 0.02 = small, 0.13 = medium, 0.26 = large for η^2 , and .01 = small, .06 = medium, .14 = large for η_p^2 (University of Cambridge, n.d.).

Firstly, in 2014, Weems et al. (2014) examined a predominantly behavioural group-based CBT approach. This longitudinal study measured TA levels at pre-intervention, post-intervention, and three follow-ups seven months, 13-18 months, and 23-29 months following pre-intervention measures. TA levels were compared within-subjects, as well as between intervention and control groups. Using hierarchical linear modelling, this study revealed a significant effect of group condition [$t(307) = 2.82, p < .01$] in relation to TA levels. The intervention group demonstrated a significant reduction [$t(164) = 8.25, p < .001, d = .84$] from pre-intervention to post-intervention with a large effect size. The waitlist control group also demonstrated a significant but smaller reduction [$t(117) = 4.73, p < .001, d = .48$] in TA levels with a medium effect size. An independent samples t-test indicated that those who received the treatment had significantly lower TA levels than the waitlist group at post-intervention [$t(281) = -2.44, p < .05, d = -.30$]. Follow-up analysis revealed significant linear ($t(447) = -6.42; p < .001$) and quadratic ($t(357) = 4.33; p < .001$) components (overall $\chi^2(247) = 504.70, p < .001$) in the intervention group. The curvilinear pattern indicated

significant reductions from pre-intervention to post-intervention. This was followed by smaller declines from post-intervention to the three follow-up points.

Later in the year 2014, a 'looking ahead' strategy was examined by Mavilidi et al. (2014). TA levels were measured post-intervention for three timepoints: before, during, and after the test. TA was then compared as a within-subjects factor, with test strategy and TA levels compared as between-subjects factors. Results displayed significant main effect of time [$F(2, 222) = 30.13, p < .001, \eta_p^2 = 0.21$] and anxiety level [$F(2, 111) = 7.87, p = .001, \eta_p^2 = 0.12$], however, the main effect of test strategy was not significant [$F(1, 111) = 3.20, p = .076, \eta_p^2 = 0.02$]. Additionally, no significant differences were found for interactions between time and TA level [$F(4, 222) = 1.22, p = .304, \eta_p^2 = 0.02$], time and test strategy [$F(2, 222) = 3.02, p = .051, \eta_p^2 = 0.03$], time, anxiety level, and test strategy, [$F(4, 222) < 1, p = .647, \eta_p^2 = 0.01$]. In relation to math test scores, the experimental group ($M = 6.14$) demonstrated significantly better performance than the control condition ($M = 5.00$) with a medium effect size ($d = 0.62$).

Following this, Carsley et al. (2015) found that state TA significantly reduced following a mindfulness-based colouring activity in children in a private school. Anxiety scores were $M = 31.769$ at pre-intervention and $M = 28.293$ at post-intervention [$t(25) = 2.925, p = .007$]. However, it is difficult to know how much of this effect is due to this mindfulness-based intervention as the free drawing/colouring control group showed similar reductions whereby $M = 28.308$ at pre-intervention and $M = 25.846$ at post-intervention [$t(25) = 3.032, p = .006$]. No significant difference was found between the control group and the mindfulness group; however, a significant gender*condition interaction with a medium effect size was found which showed that males benefited, in terms of anxiety reduction, from both groups but that females benefited from the mindful condition only [$F(1, 47) = 2.90, p = .095, \eta_p^2 = .058$]. Similar results were produced by Carsley and Heath (2019) in public school children who reported significant decreases in TA following interventions with a large effect size whereby $M = 27.20$ at pre-intervention and $M = 25.96$ at post-intervention for the mindful condition and $M = 27.92$ at pre-intervention and $M = 25.82$ at post-intervention for the free drawing condition [$F(1, 148) = 32.07, p < .001, \text{Wilk's } \lambda = .822, \eta_p^2 = .178$]. However, unlike Carsley et al. (2015), they did not find any gender differences. Carsley et al. (2015) suggested that the gender discrepancy may be attributed to the alternative way boys approached the task; they scribbled over the entire mandala rather than approaching the

intricate shapes within. However, given their contradictory findings, Carsley and Heath (2019) concluded that further research would be required to clarify gender differences.

In 2016, Yeo et al. (2016) studied an intervention based on CBT treatment strategies. A significant reduction in TA levels at follow-up ($M = 2.11$, $SD = 0.71$), two months after the intervention, compared to baseline ($M = 2.36$, $SD = 0.62$) and post-treatment ($M = 2.29$, $SD = 0.66$) was observed in participants who received the intervention. However, there were no significant changes from baseline to post-treatment. Additionally, lower levels of anxiety in the CBT group compared to the control group were reported at follow-up; the mean change in TA levels was significantly greater for the CBT group ($M = 0.26$, $SD = 0.60$) compared to the control group ($M = -0.01$, $SD = 0.41$), [$t(113) = -2.74$, $p = .007$, two-tailed]. The effect size of the differences in the mean change in TA scores was medium ($d = 0.52$).

Another study, published later in 2016, is Thompson et al.'s (2016) examination of a moderate to vigorous physical activity intervention. TA levels were measured at post-intervention and compared between groups. Results indicated no significant differences in TA levels between the intervention group and control group for either math test ($\beta = -0.25$, $SE = 1.7$, $p = 0.888$) or reading test ($\beta = 0.82$, $SE = 1.7$, $p = 0.631$). However, differences by sex and race were observed. Firstly, in a race-stratified adjusted model, amongst African American students, participants in the intervention group had lower physical state TA levels for both math (2.3 points lower, $p = 0.012$) and reading tests (2.0 points lower, $p = 0.040$), with large and medium effect sizes respectively ($d = 0.51$ and $d = 0.40$), as well as lower overall TA scores for the math test (9.7 points lower, $p = 0.018$) with a medium effect size ($d = 0.48$). Secondly, amongst male participants overall, intervention participants' TA scores for off-task behaviours were 0.8 points higher on the reading test than the control group ($p = 0.040$), although the effect size was reported to be small ($d = 0.03$). In regard to test performances, there was no statistically significant differences in the changes in scores over time for either reading ($p = 0.188$, $d = 0.10$) or maths ($p = 0.522$, $d = 0.02$) between groups.

A study of the effects of an 'Integrated Training Programme' on a female population was then carried out in 2018 by Pourtaleb et al. (2018). Considerable reductions in TA were reported from pre-intervention to post intervention for both experimental groups such that $M = 61.53$ at pre-intervention and $M = 40.86$ at post-intervention for group one, and $M = 57.93$ at pre-intervention and $M = 37.26$ at post-intervention for group two. The effect of the intervention in reducing TA was significant ($F = 19.55$, $p < 0.05$) with a large effect size (η_p^2

= 0.52). Additionally, significant mean differences between experimental groups and control groups were observed with mean differences of $M = 18.67$ between experimental group one and control group one, $M = 19.38$ between experimental group one and control group two, $M = 20.04$ between experimental group two and control group one, and $M = 20.75$ between experimental group one and control group two. Significant differences between experimental groups one and two, and between control groups one and two, were not found.

Most recently, Kurth et al. (2020) studied the effects of a mindful breathing intervention on physiological TA responses. Results were analysed by examining the effects of treatment on the binary occurrence of negative arousal peaks during the task for each time of measurement while clustering data by subject. The results revealed a significant difference between the intervention and control group ($z = 2.89, p = .004$). Despite stress level during the task acting as a significant moderator of treatment effects ($z = 1.39, p = .16$), the treatment effect remained significant ($z = 2.01, p = .035$). However, the main effect of treatment became statistically non-significant during the more difficult phase of the task ($z = 0.36, p = .72$).

2.6.9. Synthesis of Findings

This review aimed to systematically evaluate the evidence for the effectiveness of school-based TA interventions at primary-level. Following the evaluation of eight studies, it was found that the evidence for effective intervention is limited, and further research is required.

In comparison and evaluation of eight studies, the study of highest quality WoE D was Carsley & Heath's (2019) study which found that a mindfulness-based colouring intervention significantly reduced state TA in public primary schools. These findings are supported by Carsley et al.'s (2015) study who found similar results in a private school setting. These similar results across different educational settings suggest improved generalisability for the use of the intervention across these settings. It should be noted that similar results were found for the 'free drawing' control condition in both studies which suggests that free drawing is equally and significantly effective in reducing TA. The authors noted that both colouring activities could be considered mindfulness-based. Despite the high WoE D for Carsley & Heath's (2019) study, it is important to point out that this intervention was aimed at reducing participants' TA levels in their current state exclusively, and therefore, did not conduct any follow-up assessment. As a result, these mindfulness art activities may be

useful for implementation directly before a test to reduce state anxiety. However, they would not be feasible for recommendation by EPs as a long-term solution for TA as the evidence does not support lasting effects. Moreover, these studies utilised merely one measure of TA which impacts the accuracy in measuring this primary outcome.

The study with the next highest WoE D score is Weems et al.'s (2014) examination of a group-based CBT approach. Achieving a medium WoE D score, this study has demonstrated promising evidence for the effectiveness of this primarily behavioural approach in significantly reducing TA with lasting effects. It is important to note, however, that high attrition rates for follow-up assessments reduce the quality of this evidence in supporting the durability of intervention effects. Furthermore, only one measure of TA was administered resulting in low WoE A for measurements as multiple measures are considered to improve accuracy of outcomes.

Similarly, three remaining studies used merely one measure of TA, in the form of self-report measures, resulting in low WoE A for measurements (Kurth et al., 2020; Pourtaleb et al., 2018; Yeo et al., 2016). A further two studies used multiple reporters and/or measures of TA, although reliability was not reported (Mavilidi et al., 2014; Thompson et al., 2016). Consequently, the low quality of measurements across all studies could be viewed as a limitation of this review, indicating that this issue should be addressed in future research. Of these remaining five studies, two received a medium WoE D score (Pourtaleb et al., 2018; Yeo et al., 2016). Yeo et al. (2016) reported a significant reduction in TA from pre-intervention to follow-up in examining the effectiveness of a CBT-based intervention. However, no significant difference was observed from pre-intervention to post-intervention indicating a potential lag effect. Notably, due to lack of random assignment of participants to intervention and control groups in this study, it is difficult to infer causality; this means that it is not possible to determine if the effects are caused by the intervention or an alternative confounding variable (Rosenthal & Rosenthal, 2011). As a result of this methodological weakness and potential lag effect, this intervention could not be considered by EPs for recommendation of use. Pourtaleb et al. (2018) then studied the effects of an Integrated Training Programme on TA and found significant reductions in TA post-intervention. However, a lack of follow-up assessment means the durability of these effects have not been determined. Furthermore, this study was limited in terms of gender balance as it focused on a female population exclusively. Therefore, there is no evidence for the generalisability of these findings to co-educational or exclusively male schools.

The final three studies included in this review received low WoE D scores (Kurth et al., 2020; Mavilidi et al., 2014; Thompson et al., 2016) which means that no inferences could be drawn with confidence from the findings.

Coinciding with the findings of previous reviews of TA intervention (Ergene, 2003; Soares & Woods, 2020; Von Der Embse et al., 2013), the key findings of the present review indicate that the evidence for the effectiveness of school-based interventions at primary-level is still limited. Similar to the conclusions of the most recent of these reviews (Soares & Woods, 2020), minimal replication of studies and substantial heterogeneity of intervention approaches in this review further highlights the inadequacy of the evidence. This is because independent replication is required for an intervention approach to be considered evidence-based (Braden & Shernoff, 2008). Only two studies (Carsley & Heath, 2019; Carsley et al., 2015) included evaluation of the same approach, a mindfulness colouring activity. However, as previously mentioned, despite high quality evidence, this intervention could not be recommended due to several drawbacks of the approach. Furthermore, these studies were both carried out by the same researchers, and therefore, are not considered independent. Replication studies carried out by independent researchers are required for objective evaluations (Rumrill et al., 2020). A further three studies examined similar approaches in that they were grounded in CBT approaches combined with study skills teaching. However, these studies varied in terms of their focus and delivery. Weems' et al. (2014) programme emphasised behavioural approaches, with the absence of cognitive restructuring, and was delivered in small group settings over five to six sessions, whereas Yeo et al. (2016) included both behavioural and cognitive modification strategies, in the form of positive self-talk, in a whole-class setting in just four sessions. Pourtaleb et al. (2018) also combined behavioural techniques with cognitive restructuring and study skills during 14 group sessions. In their approach, cognitive restructuring involved terminating negative thoughts and replacing with positive thoughts using self-admiration and positive sentences. Moreover, findings from these studies varied; participants in Weems' et al.'s (2014) and Pourtaleb et al.'s (2018) studies demonstrated significant TA reductions at post-intervention, as well as at follow-ups for Weems' et al.'s (2014) longitudinal study, whereas Yeo et al.'s (2016) sample did not demonstrate significant reductions until follow-up. The remaining studies evaluated three diverse approaches, including vigorous physical activity (Thompson et al., 2016), mindful breathing (Kurth et al., 2020), and a looking ahead strategy (Mavilidi et al., 2014), without replication. Overall, the most commonly adopted approaches included cognitive and

behavioural strategies combined with skills teaching. However, these were evaluated in just three of the eight studies and varied in terms of specific protocols. It is worth noting, however, that Weems et al.'s (2014) intervention protocol was examined in a previous study (Weems et al., 2009), not included in this review, with similar findings in relation to post-intervention outcomes. Although this suggests increased promise for the effectiveness of this programme, an independent and objective evaluation has not yet been conducted.

In conclusion, study limitations and heterogeneity of TA intervention approaches indicate that evidence for effective intervention at primary-level is restricted. In an Irish Educational Psychology context, this means that EPs cannot rely on the current research in their role of supporting schools to meet the needs of their pupils based on the best available evidence.

2.6.10. Outstanding Issues and Future Research

As discussed, the evidence for effective intervention in reducing TA at primary-level is restricted. It is, therefore, vital to address the outstanding issues of the literature in future research. Firstly, a major limitation of this review is that all included studies were conducted in countries outside of Ireland. This means that the generalisability of these findings to an Irish population cannot be assumed as school systems may operate differently in different cultures (Von Der Embse et al., 2013). Therefore, future research should address intervention in an Irish setting.

Furthermore, the strongest evidence in effectively reducing TA was a mindfulness-based colouring activity, however, it was focused solely on TA in its current state and did not conduct follow-up assessment. It is recommended, therefore, that future research should focus on intervention aimed at lasting reduction and prevention of TA and conduct follow-up assessment to examine durability of effects (Whitley et al., 2013). In consideration of the methodological weaknesses of the studies mentioned above, future research should also endeavour to ensure higher quality methodology through the use of a randomised-control design to accurately determine intervention effects (Lewis-Beck et al., 2004), measures which are reliable for measuring TA in the specific participant sample, as well as measuring from multiple sources, to ensure accuracy in measuring intervention effects (Mazurek Melnyk & Morrison-Beedy, 2018), and a gender balanced participant sample to improve generalisability of results. Finally, replicating interventions which have previously demonstrated promising evidence is required to extend the research.

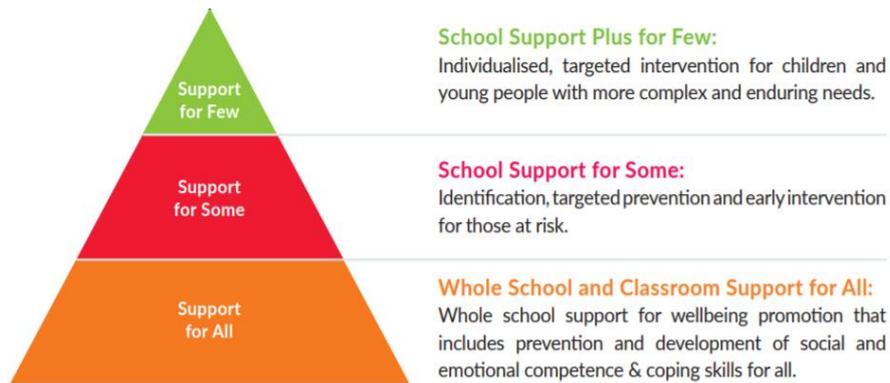
In deciding on a particular intervention for future research, it is worth looking at previous research which has shown promising evidence for behavioural strategies with the incorporation of cognitive strategies and study skills (Ergene, 2003; Von Der Embse et al., 2013). It is possible that primary-level children would also benefit from these strategies when tailored for their age group. The literature on psychological intervention for children advises replication of interventions which have already been found to be promising to advance the evidence-base for efficacious approaches, rather than developing new intervention protocols (Roberts et al., 2018). Based on the findings from this systematic review, Weems et al.'s (2014) study highlights promising evidence for a manualised programme based on the approaches outlined above and among this age group. This manualised programme (Weems, 2015) has been successfully implemented in targeted group settings with ethnic minority youth from low socioeconomic backgrounds in the USA, including in an earlier study with 13-16-year-olds (Weems et al., 2009). This earlier study also revealed significant gains in academic achievement. Previous systematic reviews (Soares & Woods, 2020; Von Der Embse et al., 2013), have recognised the manualised intervention as a high quality, detailed and thorough TA programme.

The manual for this intervention states that in addition to targeted individual and group delivery, it can also be delivered universally (Weems, 2015). However, research has not yet assessed the effectiveness of this type of delivery. In the move towards more inclusive educational practices, proactive universal interventions are now a key priority in the classroom (Dawson & Guare, 2018). In the context of universal programmes which are aimed at promoting mental wellbeing and social-emotional skills in primary-level pupils, a systematic review (Adi et al., 2007) and meta-analysis (Durlak et al., 2011) of the research concluded that positive effects were observed across all studies. Therefore, future research should assess the effectiveness of the universal delivery of Weems' (2015) programme in a whole-classroom setting. This would avoid the harmful effects of differentiating certain individuals for targeted delivery, such as stigma, exclusion, and bullying (Dawson & Guare, 2018; Weems et al., 2010). Furthermore, this type of delivery can act as a preventative measure for those children not currently experiencing TA. This is necessary given the increasing pressure and demands of testing and examinations as they progress through the education system (McDonald, 2001). Consequently, prevention is vital to equip pupils with the appropriate coping skills ahead of time for this increasing pressure (Yeo et al., 2016). In turn, based on a Continuum of Support (CoS) developed by NEPS (2007a), this class-wide

approach benefits a reduction in the potential need for support at the individual or small group level, and the subsequent time and resources required for this (Harrison et al., 2017). This CoS framework is illustrated in Figure 3 below in the context of wellbeing support.

Figure 3

NEPS Continuum of Support within the Wellbeing Policy Statement and Framework for Practice



Based on this proactive approach to wellbeing, Weems' (2015) intervention also has the potential to promote emotion regulation and academic self-efficacy as these are targeted in the intervention protocol. Therefore, it would also be useful to examine intervention effects on these factors. In consideration of emotion regulation, previous research has looked at two separate strategies, cognitive reappraisal (CR) and expressive suppression (ES). CR involves reframing an emotional stimulus to reduce its emotional impact; this emotion regulation strategy has been shown to significantly reduce negative emotions and increase positive emotions (Troy et al., 2018). Alternatively, ES means that people attempt to keep their emotional reactions from being detected by others in social interactions which can cause increased stress levels (Butler et al., 2003) and which has been linked to the worry component of TA (Schutz et al., 2004). In relation to CBT specifically, there is no research examining intervention effects on these variables in childhood anxiety or in test-anxious individuals. However, research has revealed that CBT strategies have contributed to an increased use of CR skills and reduced use of ES strategies in adult populations experiencing social anxiety (Goldin et al., 2014; Kivity et al., 2021). Therefore, it is possible that participants receiving Weems' emotion management CBT programme may also demonstrate similar improvements.

This research would also benefit from a focus on an Irish primary-level sample to improve the generalisability of the programme to this population. Finally, previous studies of Weems' (2015) school-based intervention were carried out by the programme author, along with colleagues, therefore, empirical research conducted by an independent party would allow for an unbiased and critical evaluation of the programme.

Chapter Three: Empirical Paper

The present chapter reports on the research study carried out following the conclusions drawn from the systematic review. This includes an outline of research questions, hypotheses, research methodology, results, and discussion of the findings.

3.1. Introduction

3.1.1. Key Issue- Test Anxiety at Primary-Level

Test Anxiety (TA) is a concept which has garnered the interest of researchers since the 20th century (Bögels et al., 2010). It can be described as an extreme fear of assessment or evaluation settings (Nata, 2007). Unlike generalised anxiety, which is not significantly correlated with test performance (Sarason & Palola, 1960), TA occurs whereby intense feelings of anxiety have a debilitating effect on performance (Larsen, 2017). TA can occur for reasons such as previous experience of failure (Putwain, 2008), low confidence and self-efficacy (Marcz, 2017), high self-prescribed and/or parental expectations (Eum & Rice, 2011; Kaya, 2004), and poor study and/or test-taking skills, (Naveh-Benjamin et al., 1987). TA is recognised as a multidimensional construct which can manifest in various forms (Stöber & Pekrun, 2004), including behavioural, cognitive, emotional and physiological symptoms (Kennedy, 2010; Sawka-Miller, 2011). If not managed, TA can lead to adverse effects such as panic attacks (Kennedy, 2010), depression (Akinsola & Nwajei, 2013), and school avoidance (Warner et al., 2018).

Although TA can impact individuals of all ages, there is growing concern over its prevalence in primary-level children due to standardised testing (Lobman, 2014). In Ireland, 50-75% of primary-level teachers reported that standardised testing causes extreme anxiety in pupils (Devine et al., 2020; O'Leary et al., 2019). In line with national policy targeting wellbeing in education (DES, 2018), these reports emphasise the importance of evidence-based universal interventions at primary-level.

3.1.2. Existing Intervention Research

Research literature surrounding TA intervention includes a broad range of approaches, such as, mindful colouring (Carsley et al., 2015) and breathing (Kurth et al., 2020), cognitive and behavioural strategies (Weems et al., 2014; Yeo et al., 2016), physical activity (Thompson et al., 2016), and a 'looking ahead' strategy (Mavilidi et al., 2014). Researchers who have systematically reviewed and compared various intervention protocols have

concluded that behavioural and/or cognitive strategies, are most effective in alleviating TA, particularly when combined with skills approaches (Ergene, 2003; Soares & Woods, 2020; Von Der Embse et al., 2013). These reviews, however, recognised the scarcity of TA intervention at primary-level. The persisting insufficiency of primary-level TA intervention literature was established in the systematic review carried out by the researcher in chapter two. This was attributed to non-significant outcomes and methodological weaknesses, such as, omission of follow-ups, high attrition rates, low quality in measuring outcomes, non-random assignment to group conditions, and lack of generalisability to an Irish setting. In line with previously published reviews, this review revealed promising evidence for cognitive-behavioural approaches combined with skills teaching when applied in primary-level settings. Overall, the key findings of the systematic review indicated that the evidence for primary-level TA intervention with lasting effects remains limited.

3.1.3. Aims of the Present Study

The identification of outstanding issues in chapter two was used to discern the aims of the present study. The primary aim was to determine the effectiveness of a school-based TA intervention (Weems, 2015) as a universal programme in an Irish primary-level setting. Secondary aims of this study were to examine intervention effects on academic self-efficacy and emotion regulation. As previously discussed, it was hoped that this would extend the research on school-based TA intervention at primary-level. In turn, the aim was to inform Educational Psychologists (EPs) in their role as scientist practitioners (Birch et al., 2015) and support pupil wellbeing in line with educational policy.

3.2. Research Questions and Hypotheses

Research questions and related hypotheses were established in consideration of the findings from the systematic review, as outlined in chapter two. Primary and secondary research questions and hypotheses are listed below.

3.2.1. Primary Research Question and Hypotheses

1. How effective is Weems' (2015) school-based TA intervention as a universal programme in reducing TA in a sample of Irish primary-level children?
 - Hypothesis one: There will be a statistically significant reduction in TA levels from pre-intervention to post-intervention in the intervention group.
 - Hypothesis two: There will be a statistically significant reduction in TA levels from pre-intervention to six-week follow-up in the intervention group.

3.2.2. Secondary Research Questions and Hypotheses

2. What effect does this intervention have on children's perceived self-efficacy of academic achievement?
 - Hypothesis three: There will be a statistically significant increase in perceived academic self-efficacy from pre-intervention to post-intervention in the intervention group.
3. What effect does this intervention have on children's emotion regulation skills?
 - Hypothesis four: There will be a statistically significant increase in the use of cognitive reappraisal (CR) strategies from pre-intervention to post-intervention in the intervention group.
 - Hypothesis five: There will be a statistically significant reduction in the use of expressive suppression (ES) from pre-intervention to post-intervention in the intervention group.

3.3. Method

3.3.1. Research Paradigm

The methodology employed in the present study is based on the positivist paradigm. This paradigm defines a worldview which is grounded in a scientific method of investigation (Kivunja & Kuyini, 2017). In terms of ontology, positivism assumes a sole, fixed reality which is both observable and measurable, and in relation to epistemology, it proposes the use of deductive approaches centred on objective, quantifiable methods (Sultan, 2018). In consideration of research methodology, this means utilising experimental quantitative methods (Kivunja & Kuyini, 2017). Therefore, as the purpose of this study is to determine the effectiveness of an intervention based on the size of TA reduction, the positivist paradigm is appropriate as it aims to objectively quantify effects.

3.3.2. Research Design and Procedure

The research design employed in this study was a randomised control trial (RCT). This is an experimental design which involves the random assignment of participants to different conditions to form two statistically equivalent groups (Myers & Dynarski, 2003). In the present study, which employed two fourth classes, each class was randomly allocated to either an intervention or waitlist control condition. This is known as a cluster RCT whereby groups of participants are randomly assigned to each condition. Data was collected using quantitative methods with self-report measures which will be outlined later.

Prior to the present study, a pilot study of two of these measures, along with a parent measure of TA, was also conducted to ensure reliability in the target sample. This involved administering the self-report measure in a classroom setting with a similar age group to the participant sample, specifically children from another fourth class in the same school. Parent measures were also sent home with participants prior to this and returned to the researcher on the day of data collection. Reliability of measures were then computed using statistical software, namely Statistical Package for the Social Sciences Version 26 (IBM SPSS 26). Reliability, specifically internal consistency, was measured using Cronbach's alpha and McDonald's omega whereby scores of $\geq .70$ were considered to have acceptable reliability. Results from the pilot study are outlined later under the description of measures.

Data collection then involved the administration of Weems' (2015) TA intervention by the researcher as a whole-class intervention. This was done as part of the Social, Personal, and Health Education (SPHE) curriculum. The programme was delivered in six weekly sessions, with each session lasting 40-45 minutes. The intervention group received the intervention first to analyse and compare outcomes with the waitlist control group, who received the SPHE curriculum as usual during the waiting period. Data was collected at three time-points: pre-intervention, post-intervention, and a six-week follow-up for both groups. Once the final measures had been administered at a six-week follow-up, the waitlist control group then received the intervention to ensure equal access. The researcher collected data by administering hard copies of three self-report measures and a demographic questionnaire in classroom settings. Participants were also given a parent measure of TA to take home and were instructed to return these to the researcher in the school setting. This parent measure was later omitted due to an insufficient number of returned forms for reliable data analysis.

3.3.3. Participant Sample

The target sample size for this initial study was established as 40 participants based on G*Power analysis. G*Power is a software used to calculate statistical power. The participant sample included children attending fourth class in an Irish primary-level setting. Participants were recruited from one school to control for potential differences in situational cultural environments between schools. This was considered important as school culture creates a psychosocial environment which can profoundly impact pupils (Kaplan & Owings, 2013). This target sample was chosen as standardised testing is mandatory in fourth class. Therefore, TA is more likely to occur in these children relative to those not subjected to

standardised testing. Although second class and sixth class are also subject to standardised testing, fourth class were targeted in this study as research has shown that increased levels of TA are observed in fourth to fifth class children (Lowe, 2019). Participants were recruited by inviting the school to participate in the study before seeking participant, parent, and teacher consent. Two classes with a total of 64 pupils were invited to participate; 53 of these individuals consented to participate in the study along with parent consent. On the day of initial data collection, nine of these pupils were absent from school, leaving 42 remaining participants. A further three participants were absent on the day of post-intervention data collection and were therefore excluded from data analysis. Of the remaining 39 participants included in the study, 22 took part in the intervention group and 17 were included in the control group. Demographic information across participant groups is outlined in Table 5.

Table 5*Demographic Information Across Participant Groups*

Variable	Intervention Group (n = 22)	Waitlist Control Group (n = 17)
Mean age (SD)	9.45 (.51)	9.38 (.50)
Gender	36% female 64% male	65% female 35% male
Birth country	Ireland- 77% (n = 17) Latvia- 4.5% (n = 1) Pakistan- 4.5% (n = 1) England- 4.5%(n = 1) Romania- 4.5%(n = 1) Lithuania- 4.5% (n = 1)	Ireland- 88% (n = 15) Romania- 6% (n = 1) United States of America- 6% (n = 1)
Primary language	English- 82% (n = 18) Polish- 4.5% (n = 1) Latvian- 4.5% (n = 1) Czech- 4.5% (n = 1) Urdu- 4.5 % (n = 1)	English- 100% (n = 17)
Reported learning difficulties or disabilities	Unspecified neurological condition- 4.5% (n = 1) Processing deficit- 4.5% (n = 1) None reported- (n = 20)	None reported- 100% (n = 17)

3.3.4. Ethical Considerations

In advance of this study, ethical issues were carefully considered and addressed in adherence to the Psychological Society of Ireland's (PSI; 2010) Code of Ethics. Ethical approval to carry out this study was granted by Mary Immaculate Research Ethics Committee (MIREC) in January 2021, with approval for amendments granted in August 2021. The primary ethical considerations for this study included informed consent, confidentiality, and the sensitive subject matter of this study.

Informed Consent. Following informed consent from the school, informed consent was sought from parents of participants as they were under the age of 18. This involved sending home hard copies of information sheets and consent forms to the parents of potential participants. Participant consent was also sought from those who received parental consent. This involved providing prospective participants with information sheets and consent forms. Informed consent was also requested from teachers before proceeding with this study. Participants were informed that they would have the right to withdraw at any point during the study without consequence.

Confidentiality. To ensure confidentiality of participants' data, identifying information was not requested on questionnaires. Instead, each participant was assigned a number which was stored in a password-protected excel codebook alongside their name. This codebook was required for the purpose of identifying participants' corresponding pre-intervention, post-intervention, and follow-up measures for data analysis. Names of participants were required on consent forms to ensure consent from both participants and parents before proceeding with their participation. Participants were informed that all names, and the name of the school, would not be identified in the final thesis and consent forms would be stored securely by the researcher only.

Sensitive Subject Matter. The sensitive issue addressed in this research was the issue of TA which had the potential to cause upset or worry for participants. Therefore, to ensure awareness of available supports, information sheets detailed contact information for relevant helplines and websites. Additionally, terms of confidentiality were agreed with participants and their parents on consent forms prior to the study. It was agreed that all information discussed in sessions would remain confidential with certain exceptions where confidentiality would be broken, specifically if there was a risk of harm to participants or other persons, or if information relating to a crime was disclosed. It was agreed that if this were to happen, the

researcher would disclose this information to a designated liaison person in Mary Immaculate College and in the school. Additionally, a protocol was developed for the risk of children expressing more generalised anxiety during discussion of TA. This protocol involved reporting to the teacher if any child was experiencing generalised anxiety to ensure they would receive the appropriate supports, as most schools are now equipped to deal with general anxiety under the Wellbeing Policy Framework and Statement for Practice (DES, 2018). This was agreed in advance with teachers on consent forms. Class teachers were also required to consent to inform parents if a child was observed to be experiencing generalised anxiety and would require support.

3.3.5. Intervention

Overview. The intervention which was evaluated in this study was Weems' (2015) TA intervention programme. The programme was designed for use in primary and post-primary school settings and can be delivered in five or six sessions. In the present study, six weekly sessions were carried out given the younger age range of participants and need for reiteration of programme techniques, and universal delivery which required additional time for class discussion. The intervention protocol was devised to teach pupils basic emotion management skills for coping with the symptoms and effects of TA. Sessions are focused primarily on behavioural strategies with few cognitive strategies. The programme begins with psychoeducation of anxiety in general and then more specifically, TA, followed by TA reduction strategies including exposure, relaxation techniques, self-evaluation and self-efficacy training, and study and test-taking skills (see Appendix N for programme manual). An overview of intervention sessions is outlined in Table 6 below, including details of how they were tailored for a whole-classroom setting.

Table 6

Overview of Intervention Sessions

Session	Overview of Topics Covered
One	The purpose of the research was reiterated verbally, and participants were reminded that they could withdraw at any time. Pre-intervention measures were administered. General anxiety and fear were discussed, followed by TA. A cognitive and behavioural conceptualisation of TA was presented by explaining anxiety responses as bodily reactions, talking to oneself, and actions/behaviours.

The universality of these experiences was discussed. In consideration of a universal approach and less relevance for some, it was discussed that individuals experience different degrees of anxiety and in response to different contexts. The benefits of a moderate amount of TA were also discussed in relation to optimising motivation. It was highlighted that relaxation exercises learned during the programme can be beneficial for all, regardless of current TA levels, and that they may be particularly useful in future or in alternative anxiety-provoking situations. Pupils' interests were explored to develop rapport, as well as informing relatable sports and entertainment analogies throughout the programme. The structure of the programme in relation to exposure was discussed while highlighting the importance of approach behaviour rather than avoidance.

Two In session two, a TA hierarchy was completed, as outlined in the manual (Appendix N). In the whole classroom setting, this was achieved by firstly asking pupils to complete the hierarchy worksheets individually. Items on the hierarchy were then written on a whiteboard and discussed verbally with the class to gain consensus on the least and most anxiety-provoking situations. Although, the class were all in agreement on the hierarchy, participants were informed that it was okay if they differed from the group and that they could follow their own hierarchy if preferred. In discussing the purpose of the hierarchy, the importance of approach behaviour was reviewed in more detail with discussion of gradually facing fears. A falling off a bicycle and getting back on analogy was discussed to support participants understanding; participants shared their own experiences of this to support normalisation of these experiences. Participants then learned relaxation techniques, including muscle relaxation, paced breathing, and relaxing imagery.

Three At the beginning of the third session, the benefits of relaxation exercises were reviewed, and these were practised with the class. Participants were encouraged to ask questions around the use of these and any difficulties they had implementing them. Ideas for modifying and memorising relaxation techniques were also covered, including 'Stop, Drop and Roll' and 'Sneaky Muscles', as

outlined in the programme manual. In the context of universal delivery and consideration of how individuals may have different preferences for certain techniques, participants were encouraged to speak about how they combined or modified relaxation approaches in a class discussion. Participants were praised for recognising their individual preferences and encouraged to practise different combinations to figure out what works best for them as individuals. The idea of practising these while facing fears was reviewed and participants repeated these during imagined exposure to the least-anxiety provoking situations on the hierarchy, for example, preparing for a test one week before, working on skills in class a week before the test, and planning a study schedule. Test-taking and study skills were also learned; participants were encouraged to come up with ideas for this as part of a class discussion. Participants were informed of additional strategies from the researcher.

Four In this session, participants continued gradual imagined exposure to items on the TA hierarchy. They practised relaxation techniques while imagining exposure to several increasingly anxiety-provoking situations, such as, the night before the test and eating breakfast on the morning of the test. Study and test-taking skills were also reviewed. Participants were given handouts with a script for relaxation exercises and list of study and test-taking tips. They were encouraged to review and practise these at home to support consolidation of skills. The concepts of self-evaluation and self-efficacy were introduced with concrete examples based on participants' interests and hobbies. Participants then drew a self-portrait illustrating themselves as calm and successful in a test. They also made lists of ideas of how they may reward themselves. Participants were then asked to share their ideas as part of a whole-class discussion.

Five At the beginning of this session, relaxation techniques were reviewed and practised. Participants then practised relaxation techniques during a fractions test in class; this was the final stage of exposure on the hierarchy. A group discussion was then carried out to review how individuals applied relaxation skills, their usefulness during the test, and explore any barriers. Self-evaluation and self-

efficacy were then reviewed, and pupils were asked to discuss in pairs how they might evaluate and reward themselves following this class test.

Six In the final session, all topics covered during previous sessions were reviewed. Pupils demonstrated appropriate knowledge of skills learned and were praised for their engagement and progress. Pupils were asked about ideas for continued progress; they discussed several suggestions, such as, listing relaxation techniques in the back of their test copies, displaying them on their bedroom walls, and regularly practising skills.

Theory and Evidence-Base. The programme content is derived from the empirical literature on TA and evidence-based intervention for childhood anxiety conditions (Weems et al., 2010). Firstly, evidence-based cognitive-behavioural strategies for anxiety symptoms in school-aged children were drawn on, specifically exposure and relaxation training (Silverman et al., 1999). Additionally, and of more relevance, the programme was influenced by evidence that these techniques are also effective in TA intervention with elementary-level pupils (Cheek et al., 2002). Based on findings from previous research (Hobson, 1996), Cheek et al.'s (2002) study also integrated reinforcing activities, such as art, which were shown to be effective as they “provide additional support and an element of fun” (Weems, 2015, p.7). Therefore, Weems’ (2015) intervention protocol incorporates the use of self-portraits created with art materials to encourage participants to visualise themselves “as calm and successful during a test” (Weems, 2015, p.8). As the name suggests, these cognitive-behavioural techniques are derived from cognitive-behavioural theory (Sapp, 1991; Segool et al., 2014), with a primary focus on behavioural responses. Psychoeducation is also included in the protocol as it is an integral starting point in interventions for anxiety problems (Lavell et al., 2020). Cognitive modification strategies were not included in the protocol as research has shown that such techniques can further impair test-anxious individuals by inadvertently elevating off-task thoughts (King et al., 1995; Prins et al., 1994). However, cognitive strategies, such as praise and self-evaluation and self-efficacy training, are utilised throughout the programme. This promotion of self-efficacy is reflective of the control value theory of emotions (Pekrun, 2006) and TA literature which suggests that perceived low self-efficacy is a contributor to TA (Banks, 2012). Moreover, this intervention accommodates pupils with poor study habits, through discussion of basic test-taking and study skills, therefore, it would

also appear to reflect the perspective of the deficits model of TA (Naveh-Benjamin et al., 1987). Correspondingly, the combined features of cognitive-behavioural techniques and study skills is indicative of the dual deficits model of TA (Stroud, 2013) which considers the impact of both cognitive interference and poor study skills in test-anxious individuals (Naveh-Benjamin, 1991).

The programme author has published evidence to support the effectiveness of the protocol in two empirical studies which found significant reductions in TA in samples of 13-16-year-olds (Weems et al., 2009), and 8-17-year-olds (Weems et al., 2014). Furthermore, systematic reviews of TA interventions have deemed Weems' programme to be thorough with high quality evidence (Soares & Woods, 2020; Von Der Embse et al., 2013).

Despite evidence for the effectiveness of this programme in a targeted group setting, it has not been evaluated as a universal programme, in other words, offered to all. The present study involved universal programme delivery in a whole-classroom setting for the advantageous reasons previously outlined.

Implementation Fidelity. To ensure implementation fidelity, the researcher utilised the programme manual published by Weems (2015). This included the use of an integrity checklist (see Appendix O) and the practising of strategies prior to implementation. Furthermore, the researcher drew on knowledge and practical experience of cognitive-behavioural strategies from academic studies and professional placement. Specifically, knowledge and role-play experience were gained from completion of a Master of Science in 'Applied Psychology: Mental Health and Psychological Therapies', which had a primary focus on cognitive-behavioural therapy (CBT). The researcher also has practical experience of CBT strategies from intervention casework while on professional placement as a Trainee Educational Psychologist (TEP).

3.3.6. Measures

Three self-report measures were administered to participants to measure the primary outcome of TA, as well as secondary outcomes of academic self-efficacy and emotion regulation skills. As previously mentioned, a parent measure of TA was also administered to enhance the accuracy in measuring TA as the primary outcome. Due to a low number of returned forms in the intervention group ($n = 4$) at the post-intervention stage, analysis of this data was not possible.

Children's Test Anxiety Scale. TA was measured as a primary outcome using the Children's Test Anxiety Scale (CTAS; Wren & Benson, 2004). The CTAS (Appendix P) is a 30-item self-report scale, with a Likert response format ranging from 'almost never = 1' to 'almost always = 4' scale. It provides an overall TA score, as well as three individual subscales: 'Thoughts', 'Autonomic reactions' and 'Off-task Behaviours'. Given the relatively small sample size of this study, CTAS subscale scores were not included in data analysis as increased comparisons would result in reduced statistical power.

As TA is the primary outcome in this study, a pilot study of this measure was conducted with a sample of seven participants to ensure reliability in the target population. Results indicated excellent internal consistency with a Cronbach's alpha (α) of .96. For the present study, reliability was also found to be in the excellent range for pre-intervention ($\alpha = .94$) and post-intervention data ($\alpha = .96$).

Perceived Self-Efficacy. As the intervention programme aims to enhance self-efficacy, which is well-established as both a causal and consequent factor of TA, participants' perceived academic self-efficacy was also measured. This was done using the 'Academic Achievement' subscale of Bandura's (2006) Children's Self-Efficacy Scale (CSES-AA). This is a self-report measure with a Likert scale of 0-100 with '0 = cannot do at all' and '100 = highly certain can do'. Subscale items were adapted for use with participants in an Irish primary-level setting and were included in the pilot study of measures. The adapted subscale (Appendix Q) contained six items based on subjects from the Irish primary curriculum. Given the small number of items on this scale, McDonald's omega (ω) was used to examine reliability as it does not rely on the same assumptions of Cronbach's alpha, particularly with tau-equivalence which leads to an underestimation of reliability if not met (Deng & Chan, 2017). Results from the pilot study indicated a reliability score in the good range ($\omega = .82$). Reliability in the present study was found to be slightly below the desirable range for acceptability ($\omega = .66$) for pre-intervention data. This was investigated and there were no specific items which reduced the reliability of this measure; therefore, no items were deleted. Reliability was in the good range at post-intervention ($\omega = .78$).

Emotion Regulation Questionnaire- Children and Adolescents. As the intervention programme aims to teach basic emotion management skills, changes in emotion regulation skills were also examined as a secondary outcome using the Emotional Regulation Questionnaire for Children and Adolescents (ERQ-CA; Gullone & Taffe, 2012). This is a 10-

item 5-point Likert scale (Appendix R) ranging from ‘strongly disagree = 1’ to ‘strongly agree = 5’. The ERQ-CA measures the use of two emotion regulation strategies: ‘cognitive reappraisal’ (CR) and ‘expressive suppression’ (ES). The authors have shown that the ERQ-CA has sound internal consistency in a sample of 10-18-year-olds with reliability coefficients of $\alpha = 0.83$ for the CR subscale and $\alpha = 0.75$ for the ES subscale, as well as demonstrating stability over time and construct and convergent validity. Therefore, given the robustness of this measure, a pilot study was not conducted. Reliability for the CR subscale in the present participant sample was slightly below the acceptable range for pre-intervention data ($\omega = .63$). Subscale items were examined to determine if reliability would improve if any items were removed, however, no problematic items were identified for removal. At post-intervention stage, the reliability of the scale was found to be in the good range ($\omega = .85$). Reliability of the ES subscale was in the good range for pre-intervention ($\omega = .72$) and post-intervention data ($\omega = .80$).

3.3.7. Data Analysis

Data analysis involved examining intervention effects for four outcomes: TA, academic self-efficacy, CR skills, and ES strategies using IBM SPSS 26.

Preliminary analysis of the data involved conducting skewness tests to determine the symmetry of data, along with kurtosis tests to identify any extreme values in scores relative to a normal distribution. Levene’s homogeneity of variance was also run to test equal variance in dependent variables across the sample.

A within-subjects and between-subjects approach to data analysis was then employed. Firstly, descriptive data provided mean scores for each measure across all three timepoints and in each group. For each primary and secondary outcome, a 2x2 repeated measures analysis of variance (ANOVA) was carried out, followed by a separate 3x2 ANOVA. The initial 2x2 ANOVAs were deemed necessary to determine pre-intervention to post-intervention outcomes of all participants, excluding those with missing data. The 3x2 ANOVA included all three timepoints and, therefore, eliminated all participants who did not take part in the follow-up. Therefore, these separate analyses were considered essential due to sizeable attrition at this time. The assumption of Mauchly’s sphericity was met for these analyses, except for analysis of CR outcomes which were then interpreted using Greenhouse Geisser instead.

Pairwise comparisons provided by ANOVAS were interpreted to determine intervention outcomes within-subjects across time. Time and time*group interactions were then examined to determine differences over time and between groups.

Effect sizes were calculated using Cohen's d for pairwise comparisons whereby d can be interpreted as .20 = small, .50 = medium, .80 = large (University of Cambridge, n.d.). Partial eta squared (η_p^2) was used to indicate effect sizes for ANOVAs (F-Tests), and are categorised as .01 = small, .06 = medium, .14 = large (University of Cambridge, n.d.). The effect size refers to the magnitude of the experimental effect (Sullivan & Feinn, 2012).

3.4. Results

3.4.1. Comparison of Baseline Data Between Groups

Prior to examination of intervention outcomes, pre-intervention data was analysed to determine any significant differences between groups in relation to the dependent variables. No significant differences were observed. A summary of pre-intervention data and pairwise comparison scores are outlined in Table 7 for the original participant sample, as well Table 8 for the follow-up sample.

Table 7

Pairwise Comparisons of Pre-Intervention Data Between Groups for Original Sample

Variable	Intervention Group Mean (SD)	Control Group Mean (SD)	Pairwise Comparison Scores
Test Anxiety- Total Score	20.60 (12.50)	21.41 (19.95)	$p = .881$
Self-Efficacy for Academic Achievement	419.50 (94.90)	446.24 (89.70)	$p = .387$
Emotion Regulation- CR	14.35 (2.52)	15.25 (2.17)	$p = .266$
Emotion Regulation- ES	6.70 (2.71)	6.64 (3.15)	$p = .955$

Table 8*Pairwise Comparisons of Pre-Intervention Data Between Groups for Follow-up Sample*

Variable	Intervention Group Mean (SD)	Control Group Mean (SD)	Pairwise Comparison Scores
Test Anxiety- Total Score	17.94 (9.93)	14.92 (12.71)	$p = .470$
Self-Efficacy for Academic Achievement	438.82 (88.87)	451.33 (99.25)	$p = .725$
Emotion Regulation- CR	14.56 (2.43)	15.18 (2.48)	$p = .510$
Emotion Regulation- ES	6.70 (2.71)	7.18 (3.19)	$p = .660$

3.4.2. Skewness and Kurtosis Analyses

Skewness tests were carried out to determine the symmetry of data, as well as kurtosis tests to identify any extreme values in scores relative to a normal distribution. In line with Byrne's (2010) guidelines, data was considered to be symmetrical and normally distributed if skewness was between -2 and +2, and kurtosis was between -7 and +7. Skewness and kurtosis scores were within the acceptable ranges for primary and secondary outcomes across three timepoints with the exception of skewness for total CTAS follow-up scores. Skewness and kurtosis scores are outlined in Table 9 and are coded as 'S' and 'K' respectively.

Table 9*Summary of Skewness and Kurtosis Scores*

Variable	Baseline	Post-Intervention	Follow-up
Test Anxiety- Total Score	S = 1.01 K = .72	S = 1.74 K = 2.37	S = 2.52 K = 6.88
Self-Efficacy for Academic Achievement	S = -.27 K = -.89	S = -.61 K = .08	S = -.05 K = -1.15
Emotion Regulation- CR	S = -.12 K = 1.02	S = -.77 K = 1.22	S = .99 K = 4.84
Emotion Regulation- ES	S = -.12 K = 1.02	S = .43 K = .88	S = -.23 K = -.45

Due to the distortion in skewness, the data for total CTAS follow-up scores were searched for outliers using boxplots. Two outliers, including one from each group condition, were identified and removed from the dataset to ensure that they would not harm the results. Data analyses were then re-run to determine skewness and kurtosis. Results indicated that both scores for CTAS follow-up scores were within the normal range with scores of .72 and .30 respectively.

3.4.3. Homogeneity of Variance

Levene's homogeneity of variance was also examined to determine if the distribution of scores relative to the mean scores of variables were equal across groups. As a rule, p -values should be $<.05$ to assume equal variance. However, it has been warned that a score $>.01$ should be considered acceptable and data should not be transformed if this assumption is met (Lorenzen & Anderson, 2018). Based on these rules, the equality of variance between groups was considered acceptable for all variables, as evidenced by the p -values outlined in Table 10.

Table 10

P-Values for Levene's Homogeneity of Variance

Variables	2x2 ANOVA		3x2 ANOVA		
	Pre	Post	Pre	Post	Follow-up
Test Anxiety- Total Score	.052	.027	.420	.654	.228
Self-Efficacy for Academic Achievement	.474	.866	.780	.075	.706
Emotion Regulation- CR	.425	.451	.808	.271	.563
Emotion Regulation- ES	.505	.697	.347	.625	.527

3.4.4. Intervention Effects on Test Anxiety

The primary aim of this study was to determine intervention effects on TA levels, as measured by the CTAS. Due to missing data, two of the 22 participants from the intervention group were excluded from data analysis. There was no missing data within the control group.

Post-Intervention Test Anxiety Outcomes. A 2x2 repeated measures ANOVA was run to examine differences in total CTAS scores from pre-intervention to post-intervention and between groups. A significant reduction in the mean of total CTAS scores was observed from pre-intervention to post-intervention ($p = .003$) in the original participant sample in the intervention group with a medium effect size ($d = 0.55$), therefore, confirming hypothesis one. The control group demonstrated no significant change in mean TA levels from pre-intervention to post-intervention ($p = .204$) with a negligible effect size ($d = 0.13$). In simpler terms, the intervention group demonstrated a 32% reduction in mean TA, while the control group decreased by only 12% over time. In examining intervention effects on the intervention group further, it was found that those with the highest level of TA at pre-intervention ($n = 7$) demonstrated a 34% decrease in mean TA scores, while those in the middle range ($n = 6$) were observed to have a mean reduction of 32%, and those with the lowest pre-intervention TA scores ($n = 7$) demonstrated a 27% decrease in mean TA levels at post-intervention.

Results from the ANOVA indicated a significant effect of time ($F = 11.16, p = .002$) with a large effect size ($\eta_p^2 = .24$). This means that TA levels significantly reduced in the overall participant sample from pre-intervention to post-intervention. However, there was no significant time*group interaction ($F = 2.07, p = .159$), despite a medium effect size ($\eta_p^2 = .06$). This lack of significance indicates that the mean reduction in TA levels from pre-intervention to post-intervention did not differ significantly between the groups. A table of means for total CTAS scores are outlined in Table 11.

Table 11

Summary of Means and Standard Deviations for Pre-Intervention and Post Intervention Test Anxiety Scores on the Children's Test Anxiety Scale (CTAS)

Outcomes	Intervention Group		Waitlist Control Group	
	Pre	Post	Pre	Post
Total CTAS Score	20.60 (12.50)	13.95 (11.68)	21.41 (19.95)	18.76 (21.45)

Intervention Effects on Test Anxiety for Follow-Up Sample. Of the original 37 participants included in pre-intervention and post-intervention CTAS data, 81% ($n = 30$) were included in data analysis for the six-week follow-up. A 3x2 repeated measures ANOVA was run to examine time and group interactions for all three timepoints to include only the

participants who took part in the follow-up ($n = 30$). This included 90% of the intervention group ($n = 18$) and 71% of the control group ($n = 12$).

In line with the 2x2 ANOVA for the original sample, pairwise comparisons of data from the follow-up sample on the 3x2 ANOVA revealed a significant reduction in CTAS scores from pre-intervention to post-intervention ($p = .003$) and a medium effect size ($d = 0.58$) in the intervention group. Furthermore, a significant reduction ($p < .001$) from pre-intervention to six-week follow-up, with a large effect size ($d = 1.23$) was observed. This result demonstrates maintenance of intervention effects, therefore, confirming hypothesis two. Additionally, a significant reduction from post-intervention to follow-up revealed further significant reductions in the intervention group ($p = .048$) with a medium effect size ($d = 0.50$). This indicates that intervention effects on TA levels are not only maintained at this time but that further gains emerged following programme completion. In contrast with findings from the 2x2 ANOVA, pairwise comparisons from the 3x2 ANOVA revealed significant reductions and medium effect sizes in mean CTAS scores from pre-intervention to post-intervention ($p = 0.038$, $d = 0.50$) and from pre-intervention to follow-up in the control group ($p = .016$, $d = 0.59$). That being said, the accuracy of results for follow-up analysis in the control group may be negatively impacted by a high attrition rate of 29%. This is based on a rule of thumb which suggests that a rate greater than 20% indicates attrition bias and poses a serious threat to validity of outcomes (Schulz & Grimes, 2002). This is particularly noteworthy when observing the considerable difference in mean pre-intervention scores for the original sample in the control group ($n = 17$; $M = 21.41$) relative to those included in the follow-up ($n = 12$; $M = 14.91$). There was no significant change from post-intervention to follow-up in the control group ($p = 1.000$, $d = 0.14$).

Finally, similar to results from pre-intervention and post-intervention data analysis, a significant main effect for time was observed ($F = 20.85$, $p < .001$) with a large effect size ($\eta_p^2 = .427$). There was no significant time*group interaction with a small effect size ($F = 1.06$, $p = .354$, $\eta_p^2 = .04$). A summary of mean CTAS scores for the follow-up sample are outlined in Table 12.

Table 12

Summary of Means and Standard Deviations for Pre-Intervention, Post Intervention and Follow-Up Test Anxiety Scores on the Children's Test Anxiety Scale (CTAS) for Follow-Up Sample

Outcome	Intervention Group			Control Group		
	Pre	Post	Follow-up	Pre	Post	Follow-up
Total CTAS	17.94	12.17	8.22 (5.13)	14.91	9.83	8.67 (7.79)
Score	(9.93)	(9.94)		(12.71)	(9.14)	

3.4.5. Intervention Effects on Academic Self-Efficacy

A secondary aim of this study was to determine intervention effects on participants' academic self-efficacy. Due to missing data, 37 of the original 39 participants were included in pre-intervention and post-intervention data for scores on the CSES-AA with 20 participants in the intervention group and 17 participants in the control group. Of these 37 participants, 78% were included in follow-up analysis ($n = 29$) with 85% of the intervention group ($n = 17$) and 71% of the control group ($n = 12$).

Firstly, a 2x2 repeated measures ANOVA was conducted to compare CSES-AA scores within and between subjects. Pairwise comparisons revealed no significant improvements in academic self-efficacy and negligible to small effect sizes from pre-intervention to post-intervention ($p = .240$, $d = 0.17$) in the intervention group. Moreover, a 3x2 ANOVA examining data from the follow-up sample revealed no significant changes from pre-intervention to post-intervention ($p = 1.000$, $d = 0.06$), pre-intervention to follow-up ($p = 1.000$, $d = 0.17$), or from post-intervention to follow-up ($p = 1.000$, $d = 0.33$) in the intervention group. Similarly, the control group demonstrated no significant improvements in academic self-efficacy from pre-intervention to post-intervention for the original sample ($p = .451$, $d = 0.11$) and from pre-intervention to post-intervention ($p = 1.000$, $d = 0.10$), from pre-intervention to follow-up ($p = .710$, $d = 0.24$), or from post-intervention to follow-up ($p = 1.000$, $d = 0.12$) for the follow-up sample.

Moreover, no significant time ($F = .062$, $p = .804$, $\eta_p^2 = .00$) or time*group effects ($F = 1.877$, $p = .179$, $\eta_p^2 = .051$) were observed in the 2x2 ANOVA or in time ($F = .330$, $p = .720$, $\eta_p^2 = 0.012$) and time*group interaction ($F = .638$, $p = .532$, $\eta_p^2 = .02$) for the 3x2 ANOVA.

Overall, results indicate that the intervention did not produce significant effects on academic self-efficacy, therefore, rejecting hypothesis three. A summary of mean CSES-AA scores CSES-AA are outlined in Table 13 below.

Table 13

Summary of Means and Standard Deviations for Baseline, Post-Intervention and Follow-up Scores on the Children's Self-Efficacy Scale for Academic Achievement

Participant Sample	Intervention Group			Control Group		
	Pre	Post		Pre	Post	
Original Sample	419.50	403.25		446.24	457.47	
	(94.90)	(100.84)		(89.70)	(109.15)	
Follow-up Sample	Pre	Post	Follow-up	Pre	Post	Follow-up
	438.82	433.53	434.65	451.33	462.25	475.16
	(88.87)	(71.41)	(83.13)	(99.25)	(123.06)	(97.27)

3.4.6. Intervention Effects on Emotion Regulation

Cognitive Reappraisal. As a result of missing data on the CR subscale of the ERQ-CA, 20 participants were included in pre-intervention to post-intervention data analysis in the intervention group, along with 16 participants from the control group. 90% ($n = 18$) of participants from the intervention group were included in follow-up analysis, while only 69% of the original 16 participants from the control group were included ($n = 11$).

Pairwise comparisons from a 2x2 ANOVA revealed a significant reduction in CR scores from pre-intervention to post-intervention with a medium effect size ($p = .036$, $d = 0.48$) in the original sample of the intervention group, therefore, rejecting hypothesis four. There was no significant change from pre-intervention to post-intervention ($p = .112$, $d = 0.50$), from pre-intervention to follow-up ($p = .115$, $d = 0.62$), or from post-intervention to follow-up ($p = 1.000$, $d = 0.15$) for the follow-up sample. The control group demonstrated no significant changes from pre-intervention to post-intervention ($p = .308$, $d = 0.37$) for the original sample, or from pre-intervention to post-intervention ($p = .429$, $d = 0.63$), pre-

intervention to follow-up ($p = .825$, $d = 0.59$), and post-intervention to follow-up ($p = 1.000$, $d = 0.02$) for the follow-up sample.

A 2x2 ANOVA indicated a significant effect for time ($F = 4.947$, $p = .003$, $\eta_p^2 = .13$) while there was no significant time*group interaction ($F = .464$, $p = .501$, $\eta_p^2 = .01$). Results from the 3x2 ANOVA revealed no significant effect for time ($F = 3.09$, $p = .065$, $\eta_p^2 = .10$) or for time*group interaction ($F = .130$, $p = .839$, $\eta_p^2 = .01$).

Table 14

Summary of Means and Standard Deviations for Baseline, Post-Intervention and Follow-up Cognitive Reappraisal Scores on the Emotion Regulation Questionnaire for Children and Adolescents

Participant Sample	Intervention Group		Control Group			
	Pre	Post		Pre	Post	
Original Sample	14.35 (2.52)	12.35		15.25	14.18	
		(5.30)		(2.18)	(3.49)	
Follow-up Sample	Pre	Post	Follow-up	Pre	Post	Follow-up
	14.55 (2.43)	12.39	11.50	15.18	13.27	13.18
		(5.56)	(6.52)	(2.48)	(3.52)	(4.14)

Expressive Suppression. Due to missing data on the ES subscale, 20 participants were included in pre-intervention to post-intervention data analysis in the intervention group, while only 14 were included from the control group. 100% ($n = 20$) of participants from the intervention group were included in follow-up data along with 79% of the original 14 participants in the control group ($n = 11$).

Pairwise comparisons from a 2x2 ANOVA revealed a significant reduction in the 'ES' subscale of the ERQ-CA from pre-intervention to post-intervention ($p = .041$) in the original sample of the intervention group with a medium effect size ($d = 0.46$), and a reduction of 19% in scores, therefore, confirming hypothesis five. In contrast, for the follow-up sample, there were no significant reductions from pre-intervention to post-intervention ($p = .147$, $d = 0.46$) or from pre-intervention to follow-up ($p = .086$, $d = 0.52$) despite medium effect sizes. No significant change was observed from post-intervention to follow-up ($p =$

1.000, $d = 0.04$). The control group demonstrated no significant change in ES from pre-intervention to post-intervention ($p = .422$, $d = 0.17$) for the original sample, or from pre-intervention to post-intervention ($p = .429$, $d = 0.18$), from pre-intervention to follow-up ($p = 1.00$, $d = 0.03$) or from post-intervention to follow-up ($p = 1.000$, $d = 0.20$) for the follow-up sample.

The 2x2 ANOVA revealed no significant effect for time ($F = .549$, $p = .464$, $\eta_p^2 = .017$) or time*group interaction ($F = 3.953$, $p = .356$) despite a medium effect size ($\eta_p^2 = .11$). Similarly, the 3x2 repeated measures ANOVA revealed no significant effect for time ($F = 1.17$, $p = .317$, $\eta_p^2 = .04$) or for time*group interaction ($F = 1.37$, $p = .272$) despite a medium effect size ($\eta_p^2 = .06$).

Table 15

Summary of Means and Standard Deviations for Baseline, Post-Intervention and Follow-up Expressive Suppression Scores on the Emotion Regulation Questionnaire for Children and Adolescents

Participant Sample	Intervention Group		Control Group			
	Pre	Post	Pre	Post	Follow-up	
Original Sample	6.70 (2.72)	5.45	6.64	7.21		
		(2.74)	(3.15)	(3.42)		
Follow-up Sample	6.70 (2.72)	5.45	5.35 (2.52)	7.64	7.09	7.09 (2.46)
		(2.74)		(3.50)	(2.47)	

3.4.7. Additional Observations

Although this research was primarily quantitative, an observation diary was also utilised by the researcher to take note of pupils' learning and progress throughout intervention delivery, as well as other noteworthy observations.

Children in both groups engaged well with the programme. The intervention group were particularly talkative during group discussions, and at times required redirection from the researcher. In the early stages of the intervention, when emphasising the importance of approach behaviour, the falling off a bike and getting straight back on analogy proved

extremely relatable in both groups with an array of individual examples of this; participants articulated their feelings around these well.

Discussion of participants' personal interests and hobbies outside of school and the application of relaxation techniques when facing evaluation in the context of these hobbies worked well in generating discussion around their usefulness. Several children noted that they had applied these strategies outside of school, for example, before dance competitions and football matches, while another child found them useful in the context of social anxiety.

After demonstrating progress in their learning of relaxation and breathing techniques, participants were observed to adapt these strategies appropriately based on their personal preferences. For example, one child stated that he only tensed the muscles in his feet during muscle relaxation, rather than all body-parts. Two popular choices were the use of the 'stop, drop, and roll' mnemonic technique and imagining a favourite place in combination with another strategy.

In the final session, participants demonstrated appropriate knowledge of concepts and strategies learned during the programme, including the various manifestations of anxiety, self-evaluation and self-efficacy, and relaxation strategies. They also discussed many ideas for maintaining progress, as previously outlined in Table 6. Discussion with the class teacher revealed that pupils enjoyed the programme and would look forward to it every week. The teacher and participants verbally discussed the usefulness of the programme in the final session and expressed gratitude at receiving access to it.

3.5. Discussion

The primary aim of this study was to determine the effectiveness of Weems' (2015) school-based TA intervention as a universal programme in an Irish primary-level setting. Secondary aims were to determine intervention effects on children's perceived academic self-efficacy and emotion regulation skills, including CR and ES.

It was hypothesised that improvements would be observed in the intervention group in each of these areas following programme completion. More specifically, it was predicted that there would be a significant decrease in TA levels and the use of ES strategies, and a significant increase in academic self-efficacy and CR skills. In order to test these hypotheses, a quantitative research design was employed in the form a cluster RCT. The programme was delivered to an intervention group of primary-level pupils by the researcher in six weekly

sessions in a whole-classroom setting. A waitlist control group was employed to compare outcomes between subjects.

Results from this study revealed mixed findings in relation to several primary and secondary outcomes. In relation to the effectiveness of the programme in reducing the primary outcome of TA levels, this study reveals promising, albeit preliminary, findings. A significant reduction in the use of ES strategies for emotion regulation was also observed. Findings indicated no significant improvements in relation to academic self-efficacy or the use of CR techniques for emotion regulation.

This section now discusses these findings and their implications, with reference to previous research literature. The strengths and limitations of this study and their potential impact on results are also discussed. Finally, recommendations to address these issues in future research are proposed.

3.5.1. Key Findings and Implications

Primary Outcome: Test Anxiety. As previously stated, the primary aim of this study was to determine the effectiveness of Weems' (2015) school-based TA intervention as a universal programme in reducing TA in an Irish primary-level context. The results from this study provide promising evidence for the programme's effectiveness in this context, although preliminary given the relatively small sample size and age range of participants.

Firstly, a significant reduction in TA was observed in the intervention group for both the original participant sample and follow-up sample following programme completion. In contrast, the original sample for the waitlist control group demonstrated no significant changes. It should be noted, however, that a significant reduction in TA was observed for the follow-up sample of the control group. That being said, there was a high attrition rate of 29% for this follow-up sample which causes threat to the accuracy of this finding (Schulz & Grimes, 2002), therefore, it should be interpreted with caution when comparing to the intervention group. Results from analysis of the intervention group's TA levels confirms the prediction that pupils who took part in the programme would demonstrate significantly reduced TA levels. This finding extends the existing literature published by the programme author, and his colleagues, which also reported significant reduction in TA levels in intervention groups (Weems et al., 2014; Weems et al., 2009). To the researcher's knowledge, this is the first study which has been conducted independent of the programme author and his colleagues. Hence, this reported outcome is pivotal in providing support for

their findings from an unbiased and critical perspective. Unlike the author's original studies, however, which reported large effect sizes for diminished TA levels in intervention groups, the present study detected a medium effect size. Of notable difference in those studies and the present study, and which may explain this difference, is the approach to delivery of the intervention. Specifically, Weems et al. (2009) and Weems et al. (2014) delivered the programme to targeted groups who were identified as having high levels of TA. Alternatively, the present study employed a universal delivery which can lead to a reduced relevance for some and consequent diminished potential for change. Accordingly, it could be reasoned that contrasting effect sizes may be explained by a diluted potential for response to intervention. Methods to address this issue in future research will be addressed in the next paragraph.

Before the significant reduction in TA levels detected in the intervention group can be confidently attributed to intervention effects, group differences must be considered. Comparison of the intervention and waitlist control groups is essential due to the potential of confounding variables contributing to reductions in TA over time. In contrast to the programme author's original studies (Weems et al., 2014; Weems et al., 2009), a comparison of TA levels from pre-intervention to post-intervention between the intervention and control groups revealed a statistically non-significant time*group interaction. There was, however, a medium effect size indicating that there was a moderate difference between the groups. Given that the likelihood of detecting significant group differences in intervention research is positively associated with a study's sample size (Sidani, 2014), one could argue that this study's relatively small sample size may have contributed to a lack of significance. This argument is supported by the detection of a moderate effect size despite this insignificance. Research literature suggests that the interpretation of effect size promotes a more scientific approach to the accumulation of knowledge as it is less affected by sample size, relative to significance tests (Sullivan & Feinn, 2012). This could suggest that, although a moderate difference between the groups was observed, statistical significance was not detected due to sample size. Accordingly, it is recommended that a larger participant sample should be considered in future research to enhance the probability of identifying statistical significance. Furthermore, given the universal approach to delivery of this intervention, a consequent diminished relevance to some participants may have resulted in weakened potential for change in TA. Subsequently, a larger sample size may also be beneficial in clarifying intervention effects by identifying those most in need of support for TA. More specifically,

and as completed in previous TA intervention research (Mok & Chan, 2016; Yeo et al., 2016), a more substantial sample size would allow researchers to categorise participants based on severity of TA. Consequently, this would help to control for this diluted relevance; this will be explained in more detail later. It is worth noting that although this approach may have led to reduced potential for change in the groups' mean scores, from an alternative perspective, it could be surmised that the significant reduction in TA levels in the intervention group adds to the evidence for programme effectiveness. That is to say, the significance of intervention effects was large enough to withstand the potential impact of lower pre-intervention scores on the overall mean change from pre-intervention to post-intervention. Furthermore, although it could be considered that this programme was less relevant for some, preliminary observations of percentage reductions in mean TA scores revealed that participants at the lower end of baseline TA scores demonstrated some reduction in TA scores, albeit at a lower level than those with higher baseline TA scores.

In relation to the durability of intervention effects on TA, a significant reduction in TA was observed from pre-intervention to six-week follow-up in the intervention group. This means that the observed improvements at the post-intervention phase were maintained for at least six weeks following programme completion. Furthermore, there was a large effect size for this significant reduction and there was an additional significant decline from post-intervention to follow-up in the intervention group. This would suggest that participants further benefited from the programme during this six-week period. This time likely gave participants the opportunity to practise and consolidate the skills and knowledge gained during the programme (Yeo et al., 2016). Although a shorter time frame than Weems et al.'s (2014) longitudinal study, this six-week follow-up provides further support for the longevity of programme effects. This is a promising discovery given the aim of this study in providing Irish children with the skills to cope with the increased demands of testing throughout their educational careers, and in line with the Wellbeing Policy Statement and Framework for Practice (DES, 2018). As this follow-up could be considered reasonably short, it is proposed that future research should include additional follow-ups to determine durability and/or further potential improvements over increased time periods, for example six months, a year, and two years.

Results from follow-up data also indicated a significant decrease in TA levels for the control group. There was no significant time*group interaction for all three timepoints, including follow-up. Although it is possible that a confounding variable, such as alternative

skills learned through the SPHE curriculum, could explain the reduction in TA levels for the overall sample, this finding cannot be confidently interpreted due to a high attrition rate in the control group which can cause serious threat to the validity of outcomes (Schulz & Grimes, 2002). Additionally, it is important to note that the mean score of TA at pre-intervention of the 12 participants who were included in follow-up analysis was notably lower than the mean score of the original 17 participants. This would suggest that the pupils who were absent at the time of follow-up were the pupils with the highest level of TA, therefore, lowering the mean of the group's follow-up scores.

Overall, these preliminary findings are promising in supporting the effectiveness of the programme, as reported in previous research studies (Weems et al., 2014; Weems et al., 2009). Consistent with the aims of the present study, this research also supports the generalisability of findings to an Irish primary-level context and provides initial evidence to support Weems' (2015) claim that the intervention is effective as a universal programme. This is comparable with previous literature which suggests that universal wellbeing and social-emotional learning programmes have a significant positive impact on primary-level pupils (Adi et al., 2007; Durlak et al., 2011). Furthermore, Weems' (2015) programme would appear to deliver more immediate effects than has been reported in previous literature on classroom-based TA intervention. Specifically, Yeo et al.'s (2016) universal CBT programme did not produce significant reductions in TA until follow-up indicating a potential lag effect. Therefore, results from the present study provide an important contribution to the research on universal programmes aimed at TA in primary-level with more expeditious effects. One plausible explanation for this difference may be the inclusion of cognitive modification strategies in Yeo et al.'s (2016) CBT study, whereas Weems' et al. (2014) omitted this approach and focused primarily on behavioural strategies. The rationale for this omission was drawn from research which demonstrates that cognitive restructuring can heighten off-task thoughts in individuals with high TA (King et al., 1995; Prins et al., 1994). Moreover, children under the age of 11-years-old often do not have adequately developed cognitive skills to understand the concept of cognitive modification (Hersen, 2005). Although Pourtaieb et al. (2018) demonstrated significant effects using a combined CBT approach, they employed an exclusively female sample, limiting the generalisability. Moreover, their study's slightly older sample of sixth-graders possess enhanced metacognitive skills, relative to 7-10-year-olds (Joyce-Beaulieu & Sulkowski, 2015). Developing cognitive skills are posited to be an important factor in children's capacity to learn the cognitive processes in CBT (Fuggle et

al., 2012). This is supported by findings from a review of CBT interventions which revealed that cognitive developmental level was a significant moderator of CBT outcomes, for children with maladaptive behaviours, with almost double the effect size reported in participants aged 11-13 relative to those aged 8-11 (Durlak et al., 1991). This suggests that behavioural strategies may be more appropriate than combined cognitive and behavioural approaches when addressing psychological intervention in younger children. The present research provides important insight into the promising efficacy of a primarily behavioural approach when addressing TA in this population. Based on the researcher's observations during group discussions, this approach may also have implications for psychological intervention in supporting feelings of anxiety outside the context of school testing. That being said, further research is required to corroborate or reject these findings, in relation to TA and broader contexts, due to study limitations, the contradictory findings from Pourtaleb et al.'s (2018) study and the significant reduction in TA in the follow-up control group sample.

Secondary Outcome: Academic Self-Efficacy. A secondary aim of this research was to evaluate the effect of Weems' (2015) programme on academic self-efficacy. Considering that self-efficacy is highly documented as a causal and consequent factor in TA, and is addressed in Weems' (2015) programme, it was predicted that participants would demonstrate significant elevations in academic self-efficacy following programme completion. Contrary to this hypothesis, no significant changes were detected in academic self-efficacy following programme completion for either group. This finding was surprising given the observed decrease in TA and the well-established link between academic self-efficacy and TA. Moreover, participants appeared to respond well to this component of the programme and demonstrated their learning through oral discussion. Conflicting with the literature, this finding would suggest that any improvements in TA are not attributed to a change in perceived academic self-efficacy. It is possible that the self-efficacy training included in the intervention programme was insufficient in meeting its goal. This may be explained by the lesser focus on self-efficacy, relative to anxiety management. That being said, it would be premature to dismiss the effectiveness of this strategy given the questionable reliability of the CSES-AA at pre-intervention phase. Furthermore, given the multi-modal nature of the programme, it is not clear which elements of the programme were most or least effective in either building self-efficacy or reducing TA. This is an area to consider in future research to identify and optimise the most effective features of the programme.

Secondary Outcome: Emotion Regulation. An additional secondary research question was ‘what effect does this intervention have on children’s emotion regulation?’. It was hypothesised that there would be a significant increase in the use of CR skills and a significant decrease in ES strategies. As previously discussed, CR involves reframing an emotional stimulus to reduce its emotional impact; whereas ES means that people attempt to keep their emotional reactions from being detected by others. The use of CR strategies are deemed to be more beneficial as they reduce negative emotions and increase positive emotions (Troy et al., 2018), whereas ES can cause increased stress levels (Butler et al., 2003) and is associated with increased worry in test-taking situations (Schutz et al., 2004). Intervention outcomes in relation to these individual constructs are outlined below.

Cognitive Reappraisal. In relation to scores on the CR subscale of the ERQ-CA, a significant decrease was discovered in the original sample of the intervention group. This directly contradicts the hypothesis that there would be an observed increase in CR in this group. This finding was not maintained at the six-week follow-up. Additionally, the follow-up sample did not demonstrate any significant reduction from pre-intervention to post-intervention. The findings did not reveal any further significant changes in CR strategies for emotion regulation within or between groups. In consideration of previous literature, other than studies which have targeted social anxiety in adults through CBT (Goldin et al., 2014; Kivity et al., 2021), there is an absence of relevant studies to compare this finding to in relation to TA or childhood anxiety intervention. In comparison to these limited available studies, these findings conflict with their observations of significant growth in CR efficacy. However, it is important to note that, unlike the present study, these traditional CBT approaches included cognitive restructuring strategies. Therefore, it is possible that the reported lack of significant increase in CR is attributed to the absence of cognitive modification techniques in Weems’ (2015) predominantly behavioural programme. The lesser cognitive component of this programme is primarily focused on building children’s self-efficacy with the aim of reducing and/or preventing TA. Therefore, considering that the cognitive component of this programme involved self-efficacy training, this finding fits with the absence of significant improvement in academic self-efficacy. Moreover, given that participants significantly reduced their TA levels despite a lack of significant improvement in CR strategies, this supports the sufficiency of Weems’ (2015) approach in eliciting change using predominantly behavioural methods. This approach is based off research which states that cognitive modification strategies can inadvertently contribute to additional off-task

thoughts (King et al., 1995; Prins et al., 1994). As previously stated, future research is required to determine the effectiveness of the individual components of the programme before dismissing their unique contributions.

In relation to the observed decrease in CR skills following programme completion in the original sample of the intervention group, it is possible that a confounding variable, such as knowledge of alternative emotion coping strategies learned during SPHE, may have contributed to this result. It is also worth noting that this significant change was not maintained at follow-up and a significant reduction from pre-intervention to post-intervention was not detected in the follow-up sample. Furthermore, reliability for the CR subscale of the ERQ-CA for pre-intervention data was questionable and, therefore, may have impacted the accuracy of this finding. Accordingly, additional research with more reliable measures is necessary before inferring definitive conclusions on this finding.

Expressive Suppression. A significant reduction in ES was observed in the original sample of the intervention group following programme completion, with a medium effect size. There was no significant change in the control group. A significant reduction was not detected from pre-intervention or post-intervention to six-week follow-up in the intervention group. However, the mean score at follow-up was slightly less than the mean score at post-intervention indicating that intervention effects were maintained. This comparison of post-intervention to follow-up scores is the most important for detection of maintenance of gains. It is likely that significance was not detected in this analysis as the 3x2 ANOVA contains more comparisons than the 2x2 ANOVA which can lead to reduced statistical power and likelihood of detecting significance (Hartas, 2015), as previously mentioned. This is evident from the observation of a significant reduction in ES from pre-intervention to post-intervention in the intervention group for the 2x2 ANOVA and absence of such significance in the 3x2 ANOVA, despite containing the exact same scores as the original sample was included in the follow-up. The significant reduction in ES for the original sample of the intervention group aligns with the conclusions of the previously mentioned studies (Goldin et al., 2014; Kivity et al., 2021) which found that participants receiving CBT for social anxiety demonstrated significant declines in the use of ES. In view of the literature which has established the negative impact of ES in giving rise to increased stress levels (Butler et al., 2003), this finding can be cautiously interpreted as a positive outcome. This result highlights the link between TA and ES given the significant reductions in TA also. In the context of TA intervention, this suggests that addressing emotional expression may be an important factor in

preventing and reducing TA. This finding is also fitting with previous research which demonstrated significant links between suppression and worry in a test-taking context (Schutz et al., 2004). That being said, there was no significant time*group interaction, despite a medium effect size, suggesting that the intervention and control group did not significantly differ in relation to mean change in ES scores. Similar to the conclusion of outcomes related to TA, it is possible that this is due to a relatively low sample size and universal delivery of the programme. Therefore, future research with an enhanced sample size and categorisation of ES scores would also be beneficial in further determining the significance of intervention effects. It would also be useful to explore the mediating effect of ES skills on TA outcomes to provide further insight into the link between these variables.

Summary of Key Findings. In summary of the main findings outlined above, the present study provides promising evidence for the effectiveness of Weems' (2015) school-based TA intervention as a universal programme in an Irish primary-level setting. Furthermore, this study demonstrates the potential benefits of this programme in relation to reduced ES. There were no significant benefits in relation to CR skills or academic self-efficacy.

In accordance with the primary aim of this study, these findings extend the research on effective primary-level TA intervention to inform EPs in their role as scientist practitioners (Birch et al., 2015). Although these findings are encouraging in light of significant effects and strengths of the study, they may be considered preliminary due to the limitations of the study, mixed findings between the original sample and follow-up sample, and statistically non-significant time*group effects. Consequently, future research to address these limitations would be beneficial in expanding this evidence. The strengths and limitations of this study, and recommendations for future research, are discussed in more detail in the next section.

3.5.2. Strengths and Limitations of the Present Study

Research Design. The use of a cluster RCT with an intervention group and waitlist control group signifies the robustness of the present study. The fundamental assumption of an RCT is that if there are two broadly equal groups whereby one group receives an intervention and the other does not, then the differences observed between groups are likely a result of the intervention (Connolly et al., 2017). Considering this assumption, this experimental design supported the aim of this study in identifying the effects of an intervention on specified

dependent variables. The integrity of controlling for potential differences through randomisation, which is standard practice when reporting an RCT (Kendall, 2003), is evident from the lack of significant differences between groups for baseline dependent variables. Furthermore, the importance of employing a control group was evident in this research from statistical results which revealed a statistically non-significant difference in the mean change in TA across groups, despite a significant reduction in TA in the intervention group. Consequently, this comparison between groups ensured that the findings were interpreted critically without prematurely attributing the significant findings in the intervention group to intervention effects.

Although this can be considered a robust research design, it was limited in terms of the non-active nature of the control condition. As discussed in chapter two, an alternative intervention group, as opposed to a non-intervention group, is preferable as it determines if the intervention is superior to an alternative intervention (Balakrishnan, 2014). In future research, it would be useful to employ another universal TA intervention, such as Yeo et al.'s (2016) classroom-based CBT approach, for comparison to determine the superiority of the present intervention. That being said, although a TA-specific intervention was not employed in the present study, the waitlist control group may be considered active in the context of SPHE input. As previously discussed, SPHE aims to enhance pupils' skills for managing feelings and coping with demanding situations (Government of Ireland, 1999). Therefore, continued access to the SPHE curriculum during the waiting period may also contribute to the learning of appropriate coping skills for TA. In future research, it would be important to clarify the focus of this SPHE input to consider its effects.

Follow-up. In line with the aim of this study to address the methodological weaknesses of previous research, the present study included a follow-up measure of primary and secondary outcomes. The inclusion of this follow-up is considered a study strength as it provides information on the durability of intervention effects. As previously mentioned, the extent of this longevity is limited by the relatively short time of six weeks when compared with Weem's (2014) longitudinal study which conducted three follow-ups up to two years following programme completion. Furthermore, in spite of this effort to measure durability of effects, results from analysis of follow-up data in the control group could not be confidently interpreted due to a high attrition rate which inferred threat to the validity of outcomes (Schulz & Grimes, 2002). Consequently, this limited the ability to draw comparisons between groups with the aim of discerning if outcomes were attributed to intervention effects or

potential confounding variables. It is, therefore, recommended that future research include additional follow-ups at later timepoints and make additional efforts to ensure lower attrition rates; suggestions for this will be discussed in more detail in chapter four.

Sample. The relatively small sample size included in this research could be considered a study limitation. As previously discussed, smaller sample sizes can contribute to an increased difficulty in detecting statistical significance. It also meant that it was not possible to determine intervention effects on the individual subscales of the CTAS. In explanation of this, increasing comparisons made during data analysis can reduce statistical power (Hartas, 2015). Therefore, to reliably make these comparisons would require a larger sample size to control for this (Jackson, 2019). For the same reason, mediation analysis was not possible to determine if intervention effects on TA may have been mediated by either self-efficacy or emotion regulation skills. These additional analyses would have provided further insight into the link between TA and these variables, as well as any mediating effects they may have in relation to TA outcomes. In future research, where a larger sample is available, it would be beneficial to carry out additional analyses examining outcomes in relation to CTAS subscales and the potential mediating effects of self-efficacy and emotion regulation skills on TA outcomes.

As previously mentioned, the participant sample was recruited from one school. Given the limited sample size, with only two groups, this was deemed necessary to control for potential differences in school cultures which may have impacted intervention outcomes (Kaplan & Owings, 2013). Consequently, this may be viewed as a study strength, as well as a study limitation due to reduced generalisability across school contexts (Smolkowski et al., 2013). In other words, because this limited sample controls for differences in educational environments and cultures, it is not clear if these differences have a moderating effect on response to intervention. In future research, with a larger sample size, it would be beneficial to recruit several schools to examine applicability across contexts.

An additional limitation concerning the participant sample of this study is the gender balance. Although the overall sample was balanced in relation to gender, the individual group conditions contained gender imbalances which presented inversely across the groups. The intervention group had a disproportionately larger percentage of males of 64%, whilst female representation was dominant in the control group with a percentage of 65%. As a consequence of these imbalances, it was not possible to examine gender differences in

intervention outcomes. That being said, due to the relatively small sample size, it would likely have been difficult to reliably draw comparisons regardless of the gender balance. Therefore, future research should endeavour to examine gender differences, where a larger sample size is available. This may be interesting given the established increased vulnerability of females to TA, relative to males (Putwain & Daly, 2014; Putwain, 2007), and the gender differences observed in previous TA intervention research (Carsley et al., 2015).

Measures. As the primary outcome of this study was TA, it was important to ensure reliable measures and multiple reporters to enhance accuracy in measuring this outcome. Unfortunately, due to an insufficient number of returned parent measures of TA at the post-intervention phase, it was not possible to include parent data in statistical analysis. This could be viewed as a study limitation as multiple reporters of a primary outcome are recommended to ensure accuracy in detecting intervention effects (Mazurek Melnyk & Morrison-Beedy, 2018). It is, therefore, recommended that future researchers develop alternative methods of data collection from parents to reduce the risk of attrition and support accuracy in measuring TA. Educational psychology practitioners have noted that parents are generally more compliant when completing online forms relative to paper forms (Jacob et al., 2016). Hence, an online forum for data collection may be more convenient and would reduce the likelihood of questionnaires being lost or forgotten by pupils. Potential explanations for the inadequate number of returned parent forms, and further suggestions to reduce the risk of this occurrence in future research, will be addressed in the next chapter.

Although it was unfortunate that insufficient parent questionnaires were returned for data analysis, the strengths in measuring TA are also noteworthy. Firstly, the CTAS was found to have excellent reliability in both a pilot study and the present study. This indicates the precision of this instrument in measuring the study's primary outcome (Kumar, 2010). Additionally, it is worth mentioning that children have been found to be consistently valid self-reporters of their anxiety (Weems et al., 2005). In particular, there is agreement that children are more accurate than parents or teachers in reporting internalising symptoms of anxiety (Maria, 2009; Weis, 2020). It could, therefore, be argued that the accuracy in measuring intervention effects on TA is adequate. That being said, multiple reporters are still advised in gaining supplementary information to promote an accurate picture of a child's anxiety symptoms (Pincus et al., 2019).

In relation to measures for secondary outcomes, there is evidence for both strengths and limitations. Promisingly, the ES subscale of the ERQ-CA displayed good reliability at both pre-intervention and post-intervention, therefore, demonstrating a strength in measuring intervention effects on ES. As previously noted, the questionable reliability of the CR subscale of the ERQ-CA, as well as the CSES-AA, make it difficult to infer conclusions on intervention effects on CR skills and academic self-efficacy. This will be discussed in more detail in chapter four, along with recommendations for future research.

Universal Delivery of Intervention. The universal approach to delivery of Weems' (2015) programme generated both advantages and disadvantages in the present study. As previously discussed, this is the first study to examine the effectiveness of the programme in a whole-classroom setting. This is valuable research in the move towards inclusive educational practice (Dawson & Guare, 2018). Furthermore, in research and subsequent evidence-based practice, universal delivery has the advantage of avoiding the harmful effects of targeted intervention, such as, stigma, bullying and exclusion (Dawson & Guare, 2018; Weems et al., 2010). Additionally, although this intervention may seem impertinent to some, this universal approach served to act as a preventative measure for those children not currently experiencing TA (Stallard, 2012). The skills learned are likely to be valuable to participants later in providing coping mechanisms for the inevitably increasing pressure of evaluation scenarios as they advance through the education system (McDonald, 2001). Furthermore, although greater improvements were observed in those with higher levels of TA, those with lower pre-intervention scores also demonstrated a 27% reduction in mean TA levels. This suggests that the programme was somewhat useful to those children also.

In spite of the benefits of the universal approach for participants, as previously mentioned, it may be considered a weakness in the research design. In particular, the inclusion of all participants in data analysis, including those with low levels of TA, could be seen as a limitation as it reduces the overall potential for improvement and subsequent mean change in TA levels. Therefore, in future research, where a larger sample size is available, it is recommended that researchers categorise participants based on severity of TA to control for the lack of relevance to some. To explain this further, participants with low TA are less likely to demonstrate significant reductions in TA as there is little room for improvement. On the other hand, significant decreases in TA are more likely to be observed in participants who report medium to high TA levels as there is greater potential for change. This is evident from examination of percentage reductions in TA scores in the present study which revealed

greater effects on those with higher levels of TA. This categorisation would allow researchers to determine intervention effects on this target population of highly test-anxious pupils. Although there are no published measures which explicitly state cut-off points for categories of TA in children, several studies have differentiated TA categories based on the reported scores of their participants. For example, Mok & Chan (2016) and Yeo et al. (2016) divided participants into high and low TA groups based on those who were above and below the median TA score, respectively. Yeo et al. (2016) did this by employing the CTAS which was reported to have excellent reliability, similar to the present study. Mok & Chan (2016) utilised the Test Anxiety Inventory with a population of 12-14-year-old students with high reliability also, although this is a slightly older population. These studies seem to ignore an important category, however; those with medium levels of TA. Given that moderate amounts of TA are considered to motivate individuals to study (Larsen, 2017), this is an important classification to consider. One study which has addressed three categories of TA is Cassady and Johnson's (2002) investigation into the reliability and validity of their own Cognitive Test Anxiety scale. This was achieved by categorising the 33% highest TA scorers as high in TA, the middle 33% as average, and the lowest 33% as low. Although the authors demonstrated that the scale is reliable and valid, it should be noted that it focuses exclusively on the cognitive features of TA, and therefore, excludes behavioural and physiological symptoms. Furthermore, this study focused on an adult population of undergraduate students, therefore, reliability with a child population has not been established. Despite the fact that Mavilidi et al.'s (2014) study categorised participants into low, medium, and high TA using the same measure, they did not report any adaptations to the measure for a child population or reliability with this participant sample. In consideration of the merits and limitations of these approaches, future research may benefit from employing the CTAS, which has excellent reliability in the target population, whilst employing Cassady and Johnson's method of classifying high, medium, and low TA based on the top, middle, and lowest 33 percentages of TA scores. Although there is no published subject-specific measure of TA, it would also be important to consider subject-specific self-efficacy which may mediate individuals' TA levels. As previously discussed, this could be achieved by comparing individual's self-efficacy across academic subjects with TA levels to gain insight into potential subject differences in TA (Vargios, 2007). A measure of generalised anxiety, such as the Beck Anxiety Youth Inventory, would also be beneficial to determine if high levels of TA occur in the context of severe generalised anxiety or if it is an independent issue.

Implementation Fidelity. As formerly stated, the researcher utilised Weems' (2015) programme manual and integrity checklist to support implementation fidelity. The researcher also has academic and professional experience and knowledge of CBT theory and practice. The culmination of these factors assisted intervention fidelity and can be considered a study strength. This fidelity is essential in achieving desired outcomes and making valid conclusions regarding intervention effects (Hagermoser Sanetti & Collier-Meek, 2019). Researchers have demonstrated that this self-report approach is the most accurate method of fidelity recording, relative to direct observation from others (Hagermoser Sanetti & Kratochwill, 2009). Nevertheless, multi-informant methods of fidelity recording, including self-report and direct observation, in future research could serve to increase confidence in conclusions regarding intervention outcomes (Mitchell et al., 2015).

3.5.3. Additional Recommendations for Future Research

Broader Age Range. The focus of this study was on TA intervention at primary-level with a sample of pupils attending fourth class. The study sample was constricted due to time constraints and, therefore, fourth classes were chosen as the most suitable candidates, as previously discussed. As an important future research direction to improving generalisability across age groups, it would be useful to extend the sample to include older pupils, such as, fifth and sixth classes.

Examining Demographic Differences. As discussed in chapter two, individual differences can heighten a person's vulnerability to TA, such as, personality differences (McDonald, 2001), cultural differences (Chen & Kaspar, 2004), intellectual and/or learning disabilities (Datta, 2014; Whitaker Sena et al., 2007), and as mentioned earlier in this chapter, gender differences. Due to the small sample size of the present study, it was not possible to compare differences in intervention effects in relation to these factors. Consequently, in future research with a larger sample size, it would be beneficial to examine baseline differences in TA scores, as well as differences in response to intervention, based on these demographic variables. This would be helpful in informing the evidence, and subsequent practice, of the potential moderating effects of these factors on response to intervention.

Teacher-led Delivery. This research provides preliminary evidence for the effectiveness of this programme in a classroom setting when delivered by a TEP. This finding is important in informing EPs in their role of intervention using evidence-based practice (Fallon et al., 2010). Many Irish teachers are also now adept at delivering universal anxiety

programmes using CBT, following research which has recommended the implementation of ‘FRIENDS’ programmes in Irish schools (Ruttledge et al., 2016). Therefore, considering this familiarity with such techniques, it is likely that teacher-led delivery of the programme would also be effective. Furthermore, the manualised nature of the programme, accompanied by an integrity checklist, supports the reliable replication of the programme (Braden & Shernoff, 2008). This manual is accessible and cost-effective for schools as it is freely available online. In future research, it would be helpful to examine this assumption that teacher-led delivery of Weems’ (2015) programme is effective. Based on the National Educational Psychological Service’s (NEPS) Continuum of Support (2007), this information could be valuable in supporting teachers to intervene at the classroom or group level, and subsequently reduce the potential need for direct involvement from an EP. In doing so, it would be important to consider adaptations for universal delivery, similar to those outlined in Table 6. This includes emphasising the universality of feelings of anxiety whilst also highlighting individual experiences of the different degrees of anxiety, different responses, and different contexts in which a person may experience these feelings, as well as exploring all interests and hobbies of pupils through class discussion, reaching a general consensus on the TA hierarchy whilst allowing pupils to follow their individual hierarchy if preferred, and encouraging pupils to practise individual adaptations to relaxation techniques to suit their preferences. Based on the researcher’s experiences, these class discussions, which promote the voice of all pupils, can take additional time. Therefore, it would be important for teachers to allow for the optional sixth session when planning, as was necessary in the present study.

3.6. Conclusion

To conclude this chapter, the findings of this initial study are promising in supporting the effectiveness of Weems’ (2015) programme in reducing TA and the maladaptive emotion regulation strategy of ES. Based on a six-week follow-up, these outcomes were maintained, therefore, highlighting the durability of effects. Furthermore, an increased benefit over time, in relation to reduced TA, was also observed. No significant benefits were observed in academic self-efficacy or CR skills; however, these findings must be interpreted with caution due to measures which demonstrated reliability scores below the acceptable level. Due to limitations of the present study, such as relatively low sample size, high attrition in the control group for follow-up data, lack of intervention relevance for some, and questionable reliability of two baseline measures, additional research is necessary to confirm or disconfirm the findings presented in this chapter.

Chapter Four: Critical Review and Impact Statement

The final chapter of this thesis discusses the study from a critical viewpoint, as well as a personal reflection on this process. Firstly, a critical overview of the adopted epistemological and ontological perspectives is presented. Additional strengths and limitations of the study are then discussed. Next, the implications of the findings and distinct contribution of the research are outlined. Finally, the researcher's personal reflection is summarised, followed by an impact statement.

4.1. Epistemological and Ontological Perspective

The term 'epistemology' has been defined as a key philosophical concept in social science research which considers questions regarding the theory of knowledge (Henn et al., 2005). This means that a researcher's epistemological perspective refers to their individual beliefs regarding the acquisition of knowledge (Schraw & Olafson, 2007). Ontologically, research is based on assumptions of what can be known, and epistemology aligns with these assumptions by answering how we can know (Mills et al., 2009). The epistemological and ontological perspectives of a researcher determine their selected research paradigm which is essential in choosing methodology (Klenke et al., 2016).

The theoretical perspective adopted in the present study aligns with a positivist paradigm. Positivism adopts an ontological worldview which assumes an objective reality (Stuart et al., 2015). The word 'positive' is said to characterise definitiveness and certainty (Jacobsen, 2020). In terms of epistemology, this paradigm posits that empirical science is necessary to validate knowledge using deductive approaches centred on objective, quantifiable methods (Sultan, 2018; Zimmerman & Kim, 2021). Accordingly, a key feature of a positivist paradigm is the adherence to a scientific method of investigation (Kivunja & Kuyini, 2017; Morrison, 2002). This occurs whereby positivist researchers propose cause-effect theories about phenomena which are then framed as hypotheses for testing (McNabb, 2020). Positivist research paradigms which avail of experimental designs are demanded in providing definitive evidence on the effectiveness of an intervention (Braden & Shernoff, 2008), and are arguably the most commonly adopted paradigms in educational research (English, 2006). This is likely attributed to a strong emphasis on evidence-based practices in the fields of education and educational psychology (Birch et al., 2015; Coburn et al., 2010). Therefore, in line with these procedural standards for intervention research, the positivist

paradigm was deemed most appropriate in meeting the purpose of this study, which was to evaluate the effectiveness of Weems' (2015) intervention.

In contrast with the chosen positivist paradigm, consideration of a constructivist or interpretivist paradigm were rejected as they did not align with the aims of this study. The constructivist and interpretivist paradigms assume a subjective epistemology and a relativist ontology (Kivunja & Kuyini, 2017). Unlike the positivist approach, constructivism and interpretivism advocate for qualitative study designs whereby researchers and participants work together to understand reality (Jacobsen, 2020). Although these paradigms have merit in gaining insight into lived human experiences (Kelly, 2006), they are not appropriate for testing cause-effect hypotheses (Rodwell & O'Connor, 1998). Consequently, these perspectives were not appropriate paradigms for evaluating intervention effectiveness.

Although positivism was the most appropriate paradigm for this study, it is not without limitations. As stated, the experimental design employed in this research was vital in determining the programme's effectiveness; this was a priority given the preliminary nature of this study. Despite the clear-cut superiority of this design for testing causal hypotheses (Loiselle et al., 2010), it fails to consider the factors of human behaviour which are not observable, such as thinking and feeling (Mertens, 2019). As mentioned, a constructivist or interpretivist paradigm would not be sufficient in determining the effectiveness of this programme. However, they may be useful in gaining supplementary information in relation to participants' subjective experiences of the programme. Accordingly, in amalgamation of these approaches, a pragmatic perspective is recommended for future research. A pragmatic paradigm contemplates the benefits of combining quantitative and qualitative approaches to achieve complementary results (Salkind, 2010). This paradigm was developed as a practical solution to the dispute between the opposing positions of positivism and constructivism (Kivunja & Kuyini, 2017). This approach would serve to extend the research on the effectiveness of the Weems' (2015) programme, given the limitations of the present study, as well as providing insight into subjective experiences of the programme, for example, through interviews. This would enhance the quality of this research by eliciting the voice of the child in line with the United Nation's (2009) Convention on the Rights of the Child.

4.2. Strengths & Limitations of the Study

In expansion of the strengths and limitations discussed in chapter three, the present section discusses the researcher's decision-making process in relation to these factors, as well

as additional dilemmas encountered, and potential explanations and solutions to certain limitations. The decisions made during research development were supported by research supervision and progression panels whereby research progress was presented, and feedback was provided by independent professionals.

4.2.1. Research Design

As previously outlined, the present study employed an experimental design with both an intervention group and a waitlist control group for comparison of outcomes. In line with what is considered to be the gold standard for quantitative research (Morgan et al., 2007), the decision was made to randomly assign these groups to either the intervention or waitlist control condition. This randomisation is considered essential in reducing the risk of differences between groups (Myers & Dynarski, 2003). Although this may be considered a study strength, it is worth noting that the present study included merely two groups who were cluster randomised based on their classes. Therefore, it could be argued that the equivalence between groups would be no different if they had not been randomly assigned to group conditions. In future research, randomisation may be more applicable and useful with a larger sample size, as recommended when recruiting a number of schools in education research (Connolly et al., 2017).

4.2.2. Sample

The target sample size for this initial study was established as 40 participants based on G*Power analysis. Based on the need for two classes for group conditions, and average class sizes, it was hoped that this number would be higher and would further enhance statistical power. Unfortunately, despite 53 pupils consenting to participation in the study, only 42 of these participants were present for pre-intervention data collection. This was due to a high volume of school absences related to a planned protest. In future research, it may be beneficial to discuss any potential for large absences such as this and/or to phone the school ahead of time to determine presence of participants on the day of data collection. Alternatively, a decision could be made to postpone data collection where there are large numbers of absences and where time is available for this. Unfortunately, this was not possible in the present study due to time constraints.

In addition to a diminished sample size at pre-intervention phase, attrition at post-intervention and missing data further lowered the number of participants to be included in data analysis. The attrition rate was intensified at the six-week follow-up for the control

group to the point where data could not be confidently interpreted. Similar to the recommendations above, foresight on the number of participants present on the day of follow-up would be beneficial in future research to mitigate the risk of high attrition.

4.2.3. Measures

As previously mentioned, the reliability of the CTAS, used to determine intervention effects on the primary outcome of TA, was excellent and can be viewed as a study strength. During research development, the incorporation of multiple reporters was planned to enhance accuracy in measuring TA outcomes (Mazurek Melnyk & Morrison-Beedy, 2018). The purpose of this was to heighten the quality of research on TA intervention as the utilisation of single measures was identified as a limitation in previous literature. The options for multiple reporters included teachers and/or parents. From an ethical standpoint, the inclusion of teachers as reporters of participants' TA was considered inappropriate and unrealistic in this universal approach as it would increase the already extensive workload of teachers. Alternatively, the use of an adapted parent measure of TA was considered appropriate as parent reports are advised in gaining insight into a child's anxiety (Pincus et al., 2019). Moreover, this would require considerably less time than teachers as it would require completion of just one questionnaire by each individual. The responsibility of parents to complete these measures at three timepoints was outlined in both information sheets and consent forms whereby parents agreed to this prior to study participation.

As previously discussed, these parent measures were excluded from data analysis due to high attrition. Despite efforts from the class teacher who sent out multiple reminders to parents to return the measures, only four were returned in the intervention group. Reflecting on the matter of predominantly self-report measures in previous research, it is possible that this limitation may be attributed to similar issues faced by researchers. Moreover, and as mentioned in chapter three, children have been found to be consistently valid self-reporters of their own anxiety (Weems et al., 2005). This may also explain this single-method approach to data collection. In consideration of potential explanations for this attrition, with the aim of troubleshooting in advance of future research, it is worth noting that the class teacher stated that materials that are sent to parents without the school logo are often considered unimportant and not returned. Therefore, in future research, with approval from the school's board of management, it may be useful to attach school-headed letters with a reminder to complete and return these measures. Furthermore, given the universal approach and relatively

young age of participants, it is possible that some parents may have felt that this research was irrelevant to their child. In future research, it would be important to emphasise the aims of the research and purpose of parent reports, for example, by holding a parent information meeting prior to commencing research. To facilitate attendance, this should be convenient for parents (Hughes & Cavell, 2013), for example, through online meeting forums or short video messages. Finally, and as previously mentioned, the use of online forms as an alternative to paper questionnaires, may promote convenience in completing these forms (Jacob et al., 2016). As a substitute or an adjunct to parent measures, it may be helpful to seek teacher-reports of TA exclusively for pupils identified as highly test-anxious. This would have the benefit of gaining additional information on those for which the intervention is most relevant, whilst reducing the expectations of teachers given the excessive workload which would be required for individual measures of a whole-class sample. Alternatively, physiological measures, which have been found to be a reliable and objective measure of TA (Roos et al., 2021) could improve the accuracy of outcomes in future research. Finally, given the definition of TA as an extreme fear of evaluation which interferes with performance, it would also be beneficial to include a measure of academic performance in future research. This would provide further insight into intervention outcomes in relation to this conceptualisation.

In relation to secondary outcomes, it has been noted that the reliability of measures for cognitive reappraisal (CR) and academic self-efficacy were questionable for pre-intervention data. Firstly, Bandura's (2006) scale was chosen to measure academic self-efficacy as it is a flexible scale designed for adaptation, and therefore, was appropriate for adjusting to an Irish primary-level context. As the scale was adapted by the researcher, it was deemed necessary to pilot the scale. The scale was piloted with a sample of seven children in the same age range as participants in the main study. Results from this pilot revealed good reliability for the adapted scale which was subsequently included in the main study. Unfortunately, however, in the present study, the reliability of this subscale was found to be questionable when measuring pre-intervention scores, despite pilot study scores and good reliability for post-intervention data. This questionable reliability is considered a study limitation as it reduces confidence in the findings related to intervention effects on perceived academic self-efficacy. It is possible that reliability was hindered by the varying degrees of participant self-efficacy for certain subjects. For example, one might have high self-efficacy in subjects such as English but low self-efficacy in Irish. Therefore, there would be greater discrepancy in their self-efficacy scores. In future research, it may be helpful to include a

larger sample in piloting scales for academic self-efficacy before deciding on measures for main studies.

Finally, the Emotion Regulation Questionnaire for Children and Adolescents (ERQ-CA) was selected for measuring emotion regulation techniques as there was strong evidence for the robustness of this scale in terms of reliability and validity (Gullone & Taffe, 2012). Consequently, it was deemed unnecessary to pilot this scale, particularly given the fact that emotion regulation was a secondary outcome. Unfortunately, despite excellent reliability at post-intervention phase, the CR subscale was slightly below the acceptable range for pre-intervention data. Similar to the point outlined above, this below acceptable reliability draws question to the accuracy of findings in relation to this secondary outcome. In hindsight, it would have been beneficial to pilot this scale before use; this is a consideration for future research to identify any issues before administration.

4.2.4. Intervention

The choice of intervention (Weems, 2015) examined in this research could be considered a study strength. A need for high-quality evidence-based TA intervention was identified following the conclusions of the systematic review, which indicated limited evidence for TA intervention at primary-level, particularly in an Irish context. Weems' (2015) programme has been commended as a thorough TA intervention and prevention programme with high quality evidence in previous reviews (Soares & Woods, 2020; Von Der Embse et al., 2013). This was based on Weems et al.'s (2014; 2009) studies which demonstrated promising evidence in reducing TA, as well as generalised anxiety and depression in Weems et. al's (2014) study. Furthermore, analysis of longitudinal data up to two years after programme completion revealed significant durability of these effects, although there were notable attrition rates. The promising findings of this programme may be attributed to the work of the author in examining the evidence for effective intervention for general childhood anxiety symptoms when developing this programme (Weems et al., 2010). As a result, the programme incorporates evidence-based strategies, including exposure and relaxation training which are drawn from Cognitive-Behavioural Therapy (CBT; Silverman et al., 1999). Additionally, and of more relevance, programme development was shaped by evidence that these techniques are also effective in TA intervention with primary-level pupils (Cheek et al., 2002), as well as activities such as art which were shown to be effective reinforcers (Hobson, 1996). Consequently, given the rigorous development of this

programme, and the supporting studies and reviews of these studies, this intervention was deemed to be the highest quality, relative to other studies evaluated in the systematic review. It was, therefore, selected for evaluation in an Irish context.

4.2.5. Universal Delivery

As previously discussed, universal programme delivery generated both advantages and disadvantages. This approach was decided upon based on the benefits of universal programmes in educational settings, as previously outlined. Furthermore, there was a lack of empirical evidence to support the claims that the programme could be delivered universally. This decision was made in careful consideration of the potential ethical issues of universal delivery, as advised in feedback from a progression panel. It was cautioned that the discussion of TA may lead to increased awareness and subsequent feelings of TA in those not currently experiencing TA. Following this recommendation, the researcher examined the literature on relevant universal programmes to determine the potential risk of harmful effects. Of particular interest were review articles which cover an array of research studies. Firstly, a systematic review (Adi et al., 2007) of universal mental wellbeing programmes in primary-level children was identified. This review concluded that positive effects were observed across all studies, with no reported negative effects. This was further supported by a meta-analysis of school-based universal social-emotional learning programmes at primary-level which also indicated positive effects exclusively (Durlak et al., 2011). Consequently, the benefits of this approach appeared to outweigh the risks, and therefore, consolidated the view that universal delivery was appropriate. That being said, given this sensitive topic of TA, it was still essential to ensure that universal delivery in the present study did not elicit harmful effects before drawing conclusions on the observed benefits of some. Accordingly, the data for the intervention group was manually searched to determine if any participants who were lower in TA at pre-intervention experienced increased levels of TA by the end of the intervention. Fortunately, no significant increases in TA were identified in this search. This indicates that the emphasis on TA throughout the programme did not produce harmful effects. Therefore, in line with previous research, it can be concluded that the universal delivery of Weems' (2015) intervention was an overall study strength and is worth considering in future research to extend the findings.

4.2.6. Timing of the Intervention

Weem's (2015) intervention was delivered to the intervention group in six weekly sessions between October and November 2021. The waitlist control group then received the intervention between January and March 2022. At the time of research development and planning, it was hoped that the intervention group would receive the intervention between April and May 2021. The rationale for this timing was based on national policy (Department of Education and Skills [DES], 2006; 2011) which states that standardised testing must be carried out in May or June. As previously discussed, standardised testing elicits a heightened prevalence for TA, relative to classroom testing (Segool et al., 2013). This is corroborated by reports from Irish primary-level teachers that pupils experience extreme anxiety in anticipation of standardised tests (Devine et al., 2020; O'Leary et al., 2019). Therefore, it was decided that this would be an appropriate time for intervention support when preparing for these tests. It was also planned that this timeline would allow for a follow-up time of three months in September 2021. However, the proposed start date of April 2021 was not possible due to difficulty in contacting school principals and gaining consent from a school to carry out the research, likely due to the Covid-19 pandemic. This difficulty was exacerbated by the decision to conduct the research in one school. This limited the number of potential schools for inclusion given the need for a school with a population large enough to have at least two fourth classes for a comparison group. Consequently, given the remaining time available to conduct this research within the timeframe of the researcher's doctoral training, the start date was delayed until October 2021, with a reduced time-period of six-weeks for follow-up. Therefore, in addition to the universal delivery of the programme, the relevance of Weems' (2015) intervention may have been reduced due to the timing of delivery. As previously discussed, this lack of relevance can diminish the potential for response to intervention, and subsequently, the detection of intervention effects. That being said, it is important to note that for individuals with high TA, both classroom testing and standardised testing can be equally anxiety-provoking (Segool et al., 2013). Therefore, the programme would still be beneficial to those individuals. Furthermore, in a national survey exploring Covid-19 practice in primary-level education, school leaders have indicated concern in relation to heightened pupil anxiety and a negative impact on standard of achievement (Burke & Dempsey, 2020), with similar findings in children's self-reports of anxiety and school success in an international context (Zengin et al., 2021). Additionally, standardised test scores in 2021, which were conducted following previous discontinuation of testing due to Covid-19 in 2020, may have

been negatively impacted by prolonged periods of school closures (National Council for Curriculum and Assessment, 2020b). Therefore, given the elevated prevalence of anxiety levels and hindered achievement, which is an antecedent to TA, it is possible that participants' TA levels may have been heightened at the time of intervention due to Covid-19. Although it is possible that this may have increased programme relevance, it should be noted that this has not yet been researched in an Irish primary-level context and therefore definitive conclusions cannot be drawn on this. In consideration of timing, it is recommended that future research focus on intervention in the lead up to standardised tests when it would be most relevant.

4.2.7. Hawthorne Effect

Another factor to be considered as a potential issue in the present study, is the Hawthorne effect. The Hawthorne effect is a research phenomenon which occurs whereby individuals demonstrate larger improvements based on the fact that they are being observed for purpose (Roselaar et al., 2019). Participants may anticipate the correct responses in an attempt to please the researcher (Jamison, 2006). Thus, a change in behaviour may be attributed to an awareness of the study aims (McCambridge et al., 2014). In turn, this can obscure research findings in relation to dependent variables (Loiselle et al., 2010). One way to control for the Hawthorne effect is to conduct a blind experiment whereby participants are deceived as to the purpose of the study (Jackson, 2014). Although this blind approach would reduce the potential for response bias, it was not implemented in the present study for several reasons. Firstly, this deception would give rise to ethical dilemmas by compromising informed consent and respect for autonomy (Resnik, 2018). Respect for autonomy and informed consent for research activities are essential in line with the Psychological Society of Ireland's (PSI) Code of Professional Ethics (clause 1.0 and 1.3.9; PSI, 2010). Therefore, withholding information would defy these professional ethical standards for research. Furthermore, blinding of participants is difficult to achieve in behavioural interventions (Tate & Perdices, 2019). This would certainly be true of the present study given the content of the programme and self-report measures which have a clear emphasis on TA. Therefore, it would not be advised that deception be applied in future research. Instead, findings should be interpreted with caution considering the potential Hawthorn effect. In provision of an objective measure of TA to control for the Hawthorne effect, and as previously mentioned, physiological measures may also be considered.

4.3. Implications of the Research

4.3.1. Implications for Knowledge of Universal TA Intervention

The primary aim of this study was to extend the research on TA intervention at primary-level in an Irish context, and to establish the effectiveness of a universal approach. As previously discussed, the chosen intervention (Weems, 2015) is rooted in CBT with a primary emphasis on behavioural strategies. This behavioural focus stems from empirical evidence which states that cognitive modification strategies can further impede test-anxious pupils by inadvertently heightening off-task thoughts (King et al., 1995; Prins et al., 1994). There are seemingly additional alignments with Pekrun's (2006) control value theory of emotions due to the incorporation of self-efficacy promotion, as well as with the deficits model based on the promotion of study and test-taking skills. Overall, this is reflective of the dual deficits model which posits that highly test-anxious individuals are impacted by both cognitive interference and poor study skills (Naveh-Benjamin, 1991).

Although preliminary in its findings, the present research provides support for this multi-modal approach to TA intervention, with an emphasis on gradual exposure and relaxation strategies, in an Irish context. This supports the findings of previous reviews (Ergene, 2003; Von Der Embse et al., 2013) of TA intervention which have concluded that behavioural and/or cognitive strategies are most effective in alleviating TA, particularly when combined with skills approaches. It also extends the generalisability of these approaches to a primary-level setting. What is more is that this research sheds light on the potential for these approaches to be applied in a universal setting. This knowledge is worthwhile in the context of inclusive educational practices whereby universal prevention programmes are now prioritised in classroom settings (Dawson & Guare, 2018); this will be discussed in more detail later. In line with the finding that there was no significant improvement in CR, the significant reduction in TA based on primarily behavioural strategies, with the absence of cognitive restructuring techniques, highlights the potential sufficiency of this approach when drawing on CBT strategies in universal settings. This is valuable information given that children in this age range do not always have adequate cognitive skills to understand the process of cognitive modification (Hersen, 2005). Based on the varying theories of TA, it could be asserted that the effectiveness of this multi-modal programme is particularly useful in a universal setting as it caters for individuals who experience TA due to varying reasons, such as, cognitive interference, low self-efficacy, and/or, poor study or test-taking skills. Future research may benefit from interviews with participants exploring which of these

components they found most useful, or if it was the combination of strategies.

Based on the observations of this initial study in relation to the reduced use of expressive suppression (ES) strategies, this intervention also has potential to support promotion of emotion regulation. As previously stated, ES refers to the efforts of individuals to keep their emotional reactions from being detected by others in social interactions which can cause increased stress levels (Butler et al., 2003). Children may engage in ES due to a fear of negative reactions from others (Savina et al., 2021). It is likely that the emphasis of Weems' (2015) programme on the universality of anxious feelings was successful in normalising these feelings and reducing this fear. Therefore, this normalising may be an important factor both in TA intervention, as well as more generally in supporting children to express their emotions and feelings of anxiety. In future research, where interviews are employed, it would be beneficial to explore the usefulness of this normalising strategy with participants in more depth.

4.3.2. Implications for Educational Policy and Curriculum

The findings from this study align with Irish policies for supporting the wellbeing of children within an educational context, and the prioritisation of inclusive educational practice. In recent years, national policy developers have given increasing consideration to concerns regarding children's health and wellbeing, including mental health and emotional wellbeing (Child Development Initiative, 2016), in education and health settings. The Wellbeing Policy Statement and Framework for Practice (DES, 2018) states that it is the role of educators to equip children with the knowledge and skills required to cope with challenges which may influence their wellbeing. Therefore, given the impact of TA on mental health and emotional wellbeing, this policy would suggest that schools have a critical role in the prevention of TA and supporting children already experiencing it (Yew Chye, 2008). This further aligns with international recommendations from the World Health Organisation (2018) which states that mental health promotion should be mainstreamed in education to support prevention of mental health issues in later life. The national Health Service Executive (HSE; 2013) has developed a framework for 'Developing a Health Promoting School' which stresses the adoption of a whole-school approach in promoting pupil wellbeing. In line with a global move towards inclusive educational practice (Dawson & Guare, 2018), this framework signifies a need for universal wellbeing approaches to be integrated within a school's curriculum. Based on the promising findings from this study in reducing TA and ES, Weems'

(2015) intervention has the potential to support the wellbeing needs of pupils in compliance with these policies.

To deliver this programme universally would require tailoring the programme for a whole class, whilst accommodating individual differences. As in the present study, this may be achieved through group discussions to generate consensus and promote universality of experiences. This should be complemented by encouraging children to individualise their use of relaxation techniques, their TA hierarchies, their goals for self-evaluation and self-rewards, as well as to consider the different evaluation contexts in which they may apply these skills. Individual differences in relation to participants' capacity for learning and language, as well as expectations, must also be considered to ensure equal access to the programme (Craig & Stevens, 2016). For example, this may require differentiating for pupils through various methods of presentation, reduced pace, providing key vocabulary in advance, giving additional time to complete activities and providing additional handouts (National Educational Psychological Service, 2007b). It would be beneficial to decide on specific supports in collaboration with the pupil as they are the expert on how they can learn best (Exley et al., 2019). To ensure effective delivery, a universal approach also requires teacher training (Adi et al., 2007); this will be discussed in more detail in the next section on implications for practice.

As discussed, education on mental health and wellbeing is recognised as an integral component of educational curricula in national policy (DES et al., 2015). Therefore, in accordance with the commitment of these policies and recommendations, an emphasis on wellbeing has become a critical feature in the Social, Personal, and Health Education (SPHE) primary curriculum in Ireland. A core aim of the SPHE curriculum is to promote individual wellbeing through education on managing feelings and developing coping skills for demanding situations (Government of Ireland, 1999). Furthermore, the National Council for Curriculum and Assessment (NCCA; 2020a) is currently in the process of revising the primary curriculum in a draft framework which recognises a call for increased allocated time for SPHE and an enhanced general emphasis on pupil wellbeing. Therefore, given the promising findings of this TA intervention, which was delivered universally in the context of SPHE, this programme has implications for the SPHE curriculum in supporting its goal of promoting the mental health and emotional wellbeing of pupils. This fits with the Wellbeing Policy Statement and Framework for Practice's (DES, 2018) suggestion that social-emotional learning programmes can be delivered universally as part of the SPHE curriculum.

The promising findings of this preliminary research in relation to the effectiveness of Weems' (2015) intervention approach in reducing TA levels may also have implications in supporting the aims of national policy on standardised testing. As previously discussed, national policy states that it is mandatory to administer standardised testing in literacy and numeracy in second, fourth, and sixth classes (DES, 2006; 2011). Literacy and numeracy skills are essential in allowing pupils to access all areas of the curriculum (NCCA, n.d.). Therefore, the purpose of standardised testing is to evaluate pupils' learning to inform teaching and determine which pupils may require additional support (DES, 2006). This is also true of general classroom tests as they are a necessary commodity for teachers to monitor pupils' progress (Buchwald & Schwarzer, 2011). In spite of the aims of these testing procedures, theory and research suggest that the evaluative nature of these tests can lead to an inaccurate representation of learning (Siegle, 2018). This results from the high-stakes nature of these evaluation settings which contributes to heightened TA (Segool et al., 2013), and consequent impaired performance and underachievement (Rana & Mahmood, 2010). Therefore, it could be surmised that a reduction in TA, as an outcome of Weems' (2015) intervention, could also be beneficial in alleviating the debilitating effect of TA on performance. Subsequently, this would provide a more accurate depiction of pupils' literacy and numeracy skills. That being said, this is merely an assumption based on TA theory and corroborating research (Weems et al., 2009). Hence, future research which measures academic achievement as an outcome of Weems' (2015) programme in a whole-classroom setting is essential before drawing conclusions on this supposition.

4.3.3. Implications for Practice

Another primary aim of this initial study was to inform Educational Psychologists (EPs) in their role of scientist practitioners (Birch et al., 2015) in an Irish context. Although the findings are preliminary due to the limitations of the present study, this research demonstrates the potential to support EPs and schools to meet the needs of children experiencing TA at primary-level. More specifically, it has implications for informing EPs of best available evidence in their key roles of intervention, consultation, and training (Fallon et al., 2010).

Firstly, the promising potential of this universal TA intervention has implications for EPs in their role of consultation. The National Educational Psychological Service (NEPS, 2020) has adopted a consultative model of service whereby EPs focus on empowering teachers to intervene effectively to meet the needs of pupils. This means that EPs have a role

in discussing and recommending suitable evidence-based interventions based on schools' needs. In line with best practice, the Continuum of Support (CoS; NEPS, 2007a) offers a flexible framework for meeting pupils' needs, including wellbeing needs. Based on this framework, EPs may advise schools of interventions at the whole-school or classroom level, group level, or individual level. Importantly, and as previously mentioned, these recommendations must be based on the best available evidence. Figure 4 illustrates the CoS within the context of national policy on wellbeing (DES, 2018, p.14).

Figure 4

NEPS Continuum of Support within the Wellbeing Policy Statement and Framework for Practice



Based on this continuum, Weems' (2015) intervention would be beneficial in promoting pupil wellbeing at the 'support for all' level. Research evaluating universal social-emotional school-based interventions indicates that classroom-based programmes can be delivered successfully by teachers (Durlak et al., 2011). In accordance with this research, it is likely that this programme could be reliably delivered by teachers with support from EPs. As Weems' (2015) programme requires some familiarity with CBT theory, there would be implications for the role of the EP in relation to training. Many primary schools across Ireland are now acquainted with CBT methods following research which has recommended the implementation of the CBT-based 'FRIENDS' programmes in Irish schools (Ruttledge et al., 2016). Advice on how to access training on this programme has been outlined in a DES (2021) support document for continued professional development (CPD). Consequently, the level of training input required from EPs may depend on teachers' prior knowledge and experience of CBT. For example, those with limited or no familiarity with cognitive-behavioural theory may require initial training, prior to more specific training on the delivery

of the programme. On the other hand, teachers who have previously completed FRIENDS training may require Weems' (2015) programme training only. In line with government support for the promotion of wellbeing CPD in schools (DES, 2021), this training could be included in future directories for wellbeing CPD. Based on the CoS, this teacher-led approach would reduce the potential need for support at the higher levels of the continuum.

Accordingly, this would benefit both schools and their pupils. Specifically, a reduced need for intervention at the 'support for some' and/or 'support for few' levels would save schools the time and resources required for this (Harrison et al., 2017). This intervention approach is also cost-effective given the freely available manual published by Weems (2015) online. Furthermore, this universal approach reduces the likelihood of EPs needing to intervene at the individual level. Consequently, the school may utilise their allocated NEPS time for alternative priorities. Finally, this approach would benefit pupils by providing them with coping skills for emotion management and TA, whilst avoiding the potential harmful effects of withdrawing children from their classrooms for targeted support. As previously discussed, such effects may include bullying, stigma, and exclusion (Dawson & Guare, 2018; Weems et al., 2010). Future research which evaluates teacher-led delivery of the intervention is necessary to extend the evidence to confirm or disconfirm the effectiveness of this approach.

4.3.4. Implications for Future Research

Another fundamental role of the EP which informs evidence-based practice is research (Fallon et al., 2010). The findings of this initial study are encouraging in identifying a universal TA programme which has potential for use in EP practice and within the primary-level curriculum. This study demonstrates that further research on this intervention approach is worthwhile and would be beneficial in addressing the limitations of the present study.

Firstly, the adoption of a pragmatic paradigm is recommended to enhance the depth and quality of this research. In terms of quantitative methods, future research should aim to address the limitations of this study, as previously outlined. This includes the use of reliable measures and multiple reporters and/or measures for primary outcomes, additional follow-ups, broader age range of participants, increased sample size, and categorisation of TA levels for comparison of intervention effects. To determine the applicability of this intervention within the classroom level of NEPS (2007) CoS, future research should also focus on examining the effectiveness of teacher-led delivery.

From a pragmatic perspective, these considerations for quantitative research should be supplemented with qualitative research. More specifically, interviews may provide insight into the experiences of pupils with varying degrees of TA. It may also clarify the preventative nature of the programme for those with low TA, as well as learning about how pupils may utilise strategies outside the context of TA. Moreover, given the multimodal nature of Weems' (2015) programme, interviews may provide an opportunity to gain insight into the most and least useful strategies for coping with TA and/or supporting emotion regulation in order to optimise these techniques in practice.

Finally, although no significant changes in perceived academic self-efficacy were observed in the present study, future research with more reliable measures is necessary before dismissing the effectiveness of this programme in this regard. Additional piloting of Bandura's (2006) scale or consideration of alternative methods may be useful to achieve this. The same considerations should be given to the findings regarding CR skills and the implications for future research.

4.4. Distinct Contribution

In line with the aim of this preliminary study, findings provide an important contribution to the literature by extending the research on TA intervention at primary-level. As previously concluded in chapter two, the research in this age group is extremely limited. In addition to expanding the research in this age group, to the researcher's knowledge, this study is the first to examine the effectiveness of TA intervention at primary-level in Ireland. This unique feature of the present study is valuable in establishing the generalisability of findings to an Irish context. This is vital given that countries' schools systems may operate differently (Von Der Embse et al., 2013), therefore, intervention effectiveness cannot be assumed across different cultures. In consideration of the specific intervention (Weems, 2015) examined, this research also provides important evidence in supporting the author's claims that it can be delivered effectively in a universal setting. Furthermore, previous studies examining this intervention were led by the programme author and the present study appears to be the first independent evaluation of the programme. In accordance with evidential standards, findings supporting the effectiveness of an intervention must be replicated independently before it can be accepted as evidence-based (Braden & Shernoff, 2008). Consequently, this research provides an important contribution to the evidence-base by yielding an impartial evaluation of the programme.

In addition to providing an important contribution to research on intervention aimed at reducing TA levels, the present study also elicited worthwhile information in relation to programme effects on emotion regulation. Firstly, this research provides initial evidence for the usefulness of this programme in reducing the use of ES strategies. This would suggest that this programme also has the potential to encourage children to express their emotions in a healthy manner when experiencing feelings of anxiety. Although significant improvements were not detected in CR strategies, this finding posits a distinct contribution in supporting the use of primarily behavioural techniques in the reduction of TA. In other words, TA significantly reduced following this approach despite a lack of cognitive modification strategies within the programme, and subsequent lack of improvement in CR. As this was not examined in the author's previous studies, this could be considered valuable information in the context of research on intervention for childhood anxiety. However, as previously discussed, additional research with more reliable measures is necessary to confirm or disconfirm this finding. Similarly, additional research examining programme effects on academic self-efficacy with reliable measures is required before confidently drawing inferences on the findings.

4.5. Personal Reflection

The researcher's personal reflection on this research experience is presented using Gibbs' (1988) six-phase reflective cycle.

4.5.1. Description

The present study was carried out as part requirement of the researcher's doctoral training in Educational and Child Psychology. At the beginning of this doctorate, the importance of research which is of an appropriate standard for a level ten qualification was emphasised. This emphasis was taken into consideration when planning this study.

4.5.2. Feelings

Upon starting out on this research process, I felt somewhat daunted by the expectations of doctoral-level research. This was perpetuated by my uncertainty regarding a research area to pursue. Once my interest was sparked in TA and appropriate interventions, this daunting feeling was slightly alleviated and my motivation was heightened. At times, however, I doubted the direction this study would take in terms of a specific TA intervention due to limited research in this area. In another sense, however, this had a motivating effect as it reassured me of the need for research in this area. Once I had decided on Weems' (2015)

intervention and received ethical approval for the study, I felt relieved and was keen to begin the action component of this research. At this stage, I began to feel disheartened due to difficulty in getting responses from schools when inviting them to participate in the study, likely due to the Covid-19 pandemic. Furthermore, I felt stressed at times when considering the likelihood that Covid-19 may result in further school closures which would impede the possibility of delivering the programme in a classroom setting. I was also conscious that standardised testing had previously been paused due to Covid-19 and that if it was discontinued for a second time the intervention may not be as relevant for pupils, or as desired by schools. When considering the uncertainty regarding these Covid-19 related issues, I felt concerned that I may need to come up with a contingency plan for an alternative research study. This stress was often heightened in times of spikes in the prevalence of Covid-19.

Once a school agreed to the study and a pilot study of measures had been completed, the next step was to conduct this research. At this point, I was conscious when preparing to deliver the intervention of the need for implementation fidelity. I was aware that a lack of integrity could impact the results regarding programme effectiveness and therefore, adversely influence the reputation of the programme. Throughout the intervention process, I felt encouraged by the engagement and positive responses of pupils to the programme. I was hopeful regarding the effectiveness of the programme, although slightly disappointed by participant and parent attrition rates.

4.5.3. Evaluation

Although there were several stressors and uncertainties in the planning of this research due to Covid-19, this research was ultimately a successful process. In delivering the intervention, I took comfort in the use of the programme manual and integrity checklist. This was essential in supporting my confidence and the successful implementation of this programme (Hagermoser Sanetti & Collier-Meek, 2019). Furthermore, pupils engaged positively with the researcher and the programme which subsequently contributed to participant-researcher rapport and supported pupil learning.

As briefly outlined above, the challenges faced during this research process included initially choosing a topic, seeking consent from a school to carry out the research, and high attrition rates. Another challenge faced during this process was the attrition rate of parent measures which led to an inability to analyse the data.

4.5.4. Analysis

Although disheartening at times when inviting schools to partake in this study with no response or refusal to take part, I understand that this was likely attributed to the stress which many school principals were facing at the time due to Covid-19. This perspective allowed me to view this as a practical issue for schools, rather than a personal rejection. Covid-19 also likely impacted pupil absences throughout data collection and intervention delivery, along with the planned protest which impacted participant numbers at the beginning of the research, as previously mentioned.

Similarly, although it was discouraging when I could not analyse the gathered data for parent measures due to attrition rates, I realise that there may have been barriers which contributed to this. For example, there may have been an absence of buy-in due to insufficient understanding of the need for such research. Due to a lack of relevance to some, this is a common issue in school-based universal anxiety prevention programmes (Lyneham & Rapee, 2011). Further considerations to explain this attrition may include loss of paper forms, and a potential lack of importance assigned to questionnaires as they were not school-headed.

As previously mentioned, I felt uncertain at the start of this process due to the common issue among level ten researchers in choosing a thesis topic (Reid, 2005). Despite the challenges faced in this overall research process, I feel that choosing an area of personal interest, and one which I perceived as important, were integral in motivating me at every stage of this project. These are important factors in supporting researcher perseverance (Bui, 2019). The progress in pupils' learning also supported my motivation and is a credit to their hard work and engagement with the programme. In the final stages of this research project, the promising findings acted as further encouragement in writing up this thesis, particularly when considering the positive impact that they may have. Although significant improvements were not observed in participants' academic self-efficacy, I feel that participants engaged well with this component of the programme and demonstrated appropriate learning. Reflecting on this finding, I realise that less time was spent on building self-efficacy, relative to anxiety management, throughout the sessions. Consequently, this may not have been sufficient in consolidating participants' learning for substantial intervention effects.

4.5.5. Conclusion

Overall, this research experience has presented me with both challenges and rewards. It has taught me the importance of contingency planning when developing a research study to ensure high quality research, for example, planning for alternative dates for data collection if there are a high level of absences to reduce attrition rates. Similarly, it has taught me the importance of enhancing parent understanding of their role in gathering data and the preventative benefits of the programme to minimise attrition, as previously discussed. Finally, for future replication of this research, I have learned that online forms would be preferable over paper forms to facilitate parent compliance (Jacob et al., 2016).

4.5.6. Action Plan

In my future work as an EP, whereby research is a continued professional role, I hope that the skills I have learned from this research process will support my professional research expertise. In doing so, I would endeavor to gain in-depth insight using a mix of both qualitative and quantitative methods where feasible and optimising parent engagement where possible. I will also continue to seek research which I feel is valuable and personally interesting to optimise my motivation and engagement with research. Furthermore, reflecting on the limitations of this study, I would hope to enhance the quality of any research I carry out in the future by addressing these issues. A table of these limitations and considerations for future research are outlined in Appendix S.

4.6. Impact Statement

Experimental research is essential in establishing evidence regarding the effectiveness of any intervention (Braden & Shernoff, 2008). In the fields of education and educational psychology, this research is necessary in accordance with the emphasis on evidence-based practice (Birch et al., 2015; Coburn et al., 2010). The aim of this evidence-based practice is to optimise positive outcomes for individuals in need of intervention (Davidson, 2005). Therefore, based on a systematic review of the literature which revealed limited evidence for school-based TA intervention, this initial study aimed to extend the research to inform evidence-based practice. This preliminary research was considered vital given the increasing concern regarding TA in Irish primary schools (Devine et al., 2020; O'Leary et al., 2019).

Although preliminary in its findings, this research provides valuable insight into the potential of Weems' (2015) programme for effective TA intervention at primary-level. Initial research, such as this, serves an essential function in determining if more rigorous

examination of an intervention is justifiable. (Smolkowski et al., 2013). These encouraging findings indicate that additional research on this programme is worthwhile and may inspire future researchers to expand the present research. Future research, if found to support the present findings, would enhance the impact of the present research in education and on educational psychology policy and practice.

In relation to EP practice, it would be premature to recommend Weems' (2015) intervention as an evidence-based universal TA programme in Irish primary schools. However, given the limited evidence which is currently available in relation to effective TA intervention, particularly in an Irish context, it is impossible to recommend an alternative intervention with superior evidence. Therefore, if schools were to seek TA intervention in consultation with EPs, for example in the role of NEPS, this may provide an opportunity to conduct future research while implementing the programme. In other words, the programme could be delivered universally by teachers, with training and support from EPs, as part of a larger-scale programme evaluation. In turn, this would serve to support test-anxious children, whilst simultaneously conducting worthwhile research to support future evidence-based practice.

This research also highlights the importance of proactive universal approaches in supporting pupil wellbeing, with a specific focus on TA, in line with inclusive educational practice. This finding supports the aims of the Wellbeing Policy Statement and Framework for Practice (DES, 2018) and has the potential to impact curriculum development in relation to SPHE. SPHE aims to promote pupil wellbeing through education on emotion management and building coping skills (Government of Ireland, 1999). A new framework for SPHE is currently being developed due to demand for an enhanced emphasis on pupil wellbeing (NCCA, 2020a). Therefore, it is fitting to promote TA prevention and reduction as part of a modern framework for SPHE, if supported by future research. Moreover, this research may also impact this development in relation to the aim of SPHE in enhancing emotion management skills by promoting healthy emotional expression. This assumption is based on the observed significant reduction in ES strategies which are maladaptive to emotion regulation.

In conclusion, this research has provided valuable insight into a promising TA intervention for primary-level children in a national context. More generally, this study also reveals the potential of this programme to promote healthy emotion regulation skills. In an

international context, this research may also inform EPs and educational professionals of proactive universal TA intervention in accordance with inclusive educational practice.

4.7. Dissemination of Research

In order for the potential impacts of this research to be achieved, the findings must be disseminated in relevant platforms. Prior to data analysis, during the intervention delivery phase, the researcher submitted an on-demand video presentation detailing the study's aims and methodology to the PSI's annual conference as part of a research symposium titled 'Promoting Wellbeing in Schools'. To extend the impact of this research, appropriate scholarly journals will also be considered for publication of the review paper and empirical paper upon completion of this thesis. Findings were also presented at a research school in May 2022 for students in Years One, Two, and Three of the doctoral training in Educational and Child Psychology in Mary Immaculate College. This may be useful to attendees in informing their future practice as Trainee Educational Psychologists and EPs. A summary of results was also provided to the school who participated in this research. Furthermore, this thesis was shared with the programme author to inform him of the findings of this independent research. Finally, in the researcher's future role as an EP, there may be additional opportunities to present this research to relevant professionals and/or conferences to further its impact.

References

- Adi, Y., Killoran, A., Janmohamed, K., & Stewart-Brown, S. (2007). Systematic review of the effectiveness of interventions to promote mental wellbeing in children in primary education. Report 1: Universal approaches: non-violence related outcomes. *Database of Abstracts of Reviews of Effects (DARE): Quality-assessed Reviews [Internet]*.
- Akinsola, E. F., & Nwajei, A. D. (2013). Test anxiety, depression and academic performance: assessment and management using relaxation and cognitive restructuring techniques. *Psychology, 4*(06), 18.
- Allen, G. J. (2016). The behavioural treatment of test anxiety: Therapeutic innovations and emerging conceptual challenges. In M. Hersen, R. M. Eisler, & P. M. Miller (Eds.), *Progress in behaviour modification: Volume 9* (pp. 81-124). Elsevier Science.
- Allen, K. A., & Kern, M. L. (2017). *School belonging in adolescents: Theory, research and practice*. Springer Singapore.
- Alpert, R., & Haber, R. N. (1960). Anxiety in academic achievement situations. *The Journal of Abnormal and Social Psychology, 61*(2), 207.
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5®)*. American Psychiatric Publishing.
- American Psychiatric Association. (2015). *Understanding mental disorders: Your guide to DSM-5®*. American Psychiatric Publishing.
- American Psychiatric Association, Frances, A., & American Psychiatric Association Task Force on DSM-IV. (1994). *Diagnostic and statistical manual of mental disorders: DSM-IV*. American Psychiatric Association.
- American Psychological Association. (2021). *Anxiety*. American Psychological Association. <https://www.apa.org/topics/anxiety>
- Anderson, & Sauser. (1995). Measurement of test anxiety: An overview. In C. D. Spielberger & P. R. Vagg (Eds.), *Test Anxiety: Theory, Assessment, and Treatment* (pp. 15-34). Taylor & Francis.
- Antony, M. M., & Norton, P. J. (2015). *The anti-anxiety workbook: Proven strategies to overcome worry, phobias, panic, and obsessions*. Guilford Publications.
- Ashcraft, M. H. (2019). Models of math anxiety. In I. C. Mammarella, S. Caviola, & A. Dowker (Eds.), *Mathematics anxiety: What is known, and what is still missing*. Taylor & Francis.
- Bagnell, A. L. (2011). Anxiety and separation disorders. *Paediatrics in Review-Elk Grove, 32*(10), 440.

- Bala, P., & Shaafiu, K. Q. (2016). Academic achievement of secondary school students in relation to their problem solving ability and examination anxiety. *The International Journal of Indian Psychology*, 3(4), 138-154.
- Balakrishnan, N. (2014). *Methods and applications of statistics in clinical trials, volume 1: Concepts, principles, trials, and designs*. John Wiley & Sons.
- Baldwin, D. S., & Leonard, B. E. (2013). *Anxiety disorders*. Karger.
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Prentice-Hall.
- Bandura, A. (2006). Guide for constructing self-efficacy scales. In T. Urdan & F. Pajares (Eds.), *Self-efficacy beliefs of adolescents* (Vol. 5, pp. 307-337).
- Banks, J., & Smyth, E. (2015). 'Your whole life depends on it': academic stress and high-stakes testing in Ireland. *Journal of Youth Studies*, 18(5), 598-616.
<https://doi.org/10.1080/13676261.2014.992317>
- Banks, S. R. (2012). *Classroom assessment: Issues and practices, second edition*. Waveland Press.
- Baron, H., Martin, T., Proud, A., Weston, K., & Elshaw, C. (2003). Ethnic group differences and measuring cognitive ability. In C. Cooper & I. T. Robertson (Eds.), *International review of industrial and organisational psychology* (Vol. 18, pp. 191-238). Wiley.
- Beidel, D. C., & Alfano, C. A. (2011). *Child anxiety disorders: A guide to research and treatment, 2nd edition*. Taylor & Francis.
- Beidel, D. C., & Turner, S. M. (1988). Comorbidity of test anxiety and other anxiety disorders in children. *Journal of Abnormal Child Psychology*, 16(3), 275-287.
- Bin Kassim, M. A., Mohd Hanafi, S. R. B., & Hancock, D. R. (2008). Test anxiety and its consequences on academic performance among university students. In A. Columbus (Ed.), *Advances in psychology research* (pp. 75-95). Nova Science.
- Birch, S., Frederickson, N., & Miller, A. (2015). What do educational psychologists do? In T. Cline, A. Gulliford, & S. Birch (Eds.), *Educational psychology: Topics in applied psychology* (Second ed.). Routledge.
- Birturk, A., & Karagun, E. (2015). The effect of recreational activities on the elimination of state-trait anxiety of the students who will take the SBS Placement Test. *Educational Research and Reviews*, 10(7), 894-900.
- Boaler, J., & Dweck, C. (2015). *Mathematical mindsets: Unleashing students' potential through creative math, inspiring messages and innovative teaching*. Wiley.

- Bochis, L., & Sandra, F. (2018). Intervention programme to reduce the level of test anxiety in a primary school class. A pilot study. *Revista Românească pentru educație multidimensională*, 10(4), 23-31. <https://doi.org/10.18662/rrem/69>
- Boekaerts, M., & Pekrun, R. (2015). Emotions and emotion regulation in academic settings. In L. Corno & E. M. Anderman (Eds.), *Handbook of educational psychology* (pp. 76-90). Taylor & Francis.
- Braden, J. P., & Shernoff, E. S. (2008). Why the need for evidence-based interventions? In R. J. Morris & N. Mather (Eds.), *Evidence-based interventions for students with learning and behavioural challenges* (pp. 9-32). Taylor & Francis.
- Brown, E. (2021). *Your kid's ok!: Celebrate uniqueness as a superpower*. Xlibris AU.
- Buchwald, P., & Schwarzer, C. (2011). Impact of assessment on students' test anxiety. In S. Jarvela (Ed.), *Social and emotional aspects of learning* (pp. 123-130). Elsevier Science.
- Buck, R. (2016). Anxiety manipulation in school-based research. *Psychology of Education Review*, 40(2), 10-16.
- Bui, Y. N. (2019). *How to write a master's thesis*. SAGE Publications.
- Burke, J., & Dempsey, M. (2020). *Covid-19 Practice in Primary Schools in Ireland Report*.
- Bushnell, G. A., Talati, A., Wickramaratne, P. J., Gameraff, M. J., & Weissman, M. M. (2020). Trajectories of childhood anxiety disorders in two generations at high risk. *Depression and Anxiety*, 37(6), 521-531.
- Butler, E. A., Egloff, B., Wilhelm, F. H., Smith, N. C., Erickson, E. A., & Gross, J. J. (2003). The social consequences of expressive suppression. *Emotion*, 3(1), 48.
- Byrne, B. M. (2010). *Structural equation modeling with AMOS: Basic concepts, applications, and programming*. Routledge.
- Bögels, S. M., Alden, L., Beidel, D. C., Clark, L. A., Pine, D. S., Stein, M. B., & Voncken, M. (2010). Social anxiety disorder: questions and answers for the DSM-V. *Depression and Anxiety*, 27(2), 168-189.
- Carducci, B. J., Nave, C. S., Di Fabio, A., Saklofske, D. H., & Stough, C. (2020). *The wiley encyclopaedia of personality and individual differences, personality processes and individuals differences*. Wiley.
- Carr, S. (2015). *Motivation, educational policy and achievement: A critical perspective*. Taylor & Francis.

- Carsley, D., & Heath, N. L. (2019). Evaluating the effectiveness of a mindfulness colouring activity for test anxiety in children. *The Journal of Educational Research, 112*(2), 143-151.
- Carsley, D., Heath, N. L., & Fajnerova, S. (2015). Effectiveness of a classroom mindfulness colouring activity for test anxiety in children. *Journal of Applied School Psychology, 31*(3), 239-255.
- Casbarro, J. (2005). *Test anxiety and what you can do about it: A practical guide for teachers, parents, and kids*. National Professional Resources Inc.
- Cassady, J. C. (2010). *Anxiety in schools: The causes, consequences, and solutions for academic anxieties*. Peter Lang.
- Cassady, J. C., & Johnson, R. E. (2002). Cognitive test anxiety and academic performance. *Contemporary Educational Psychology, 27*(2), 270-295.
- Chapell, M. S., Blanding, Z. B., Silverstein, M. E., Takahashi, M., Newman, B., Gubi, A., & McCann, N. (2005). Test anxiety and academic performance in undergraduate and graduate students. *Journal of Educational Psychology, 97*(2), 268-274.
<https://doi.org/10.1037/0022-0663.97.2.268>
- Chapman, A., Gratz, K., Tull, M., & Keane, T. (2011). *The dialectical behaviour therapy skills workbook for anxiety: Breaking free from worry, panic, PTSD, and other anxiety symptoms*. New Harbinger Publications.
- Cheek, J. R., Bradley, L. J., Reynolds, J., & Coy, D. (2002). An intervention for helping elementary students reduce test anxiety. *Professional School Counselling, 6*(2), 162-164.
- Chen, X., & Kaspar, V. (2004). Cross-cultural research on childhood. In U. P. Gielen & J. L. Roopnarine (Eds.), *Childhood and adolescence: Cross-cultural perspectives and applications* (pp. 46-80). Praeger.
- Cheng, L., Klinger, D., Fox, J., Doe, C., Jin, Y., & Wu, J. (2014). Motivation and test anxiety in test performance across three testing contexts: The CAEL, CET, and GEPT. *TESOL Quarterly, 48*(2), 300-330. <https://doi.org/10.1002/tesq.105>
- Child Development Initiative. (2016). CDI Healthy Schools' Programme
Policy Paper
Healthy Schools - Healthy Children. In.
- Cizek, G. J., & Burg, S. S. (2006). *Addressing test anxiety in a high-stakes environment: Strategies for classrooms and schools*. SAGE Publications.

- Coburn, C. E., Stein, M. K., Baxter, J., D'Amico, L., Datnow, A., Engle, R., . . . Park, V. (2010). *Research and practice in education: Building alliances, bridging the divide*. Rowman & Littlefield Publishers.
- Columbus, A. (2008). *Advances in Psychology Research*. Nova Science.
- Connolly, P., Biggart, A., Miller, S., O'Hare, L., & Thurston, A. (2017). *Using randomised controlled trials in education*. SAGE Publications.
- Contreras-Soto, A., Véliz-Campos, M., & Véliz, L. (2019). Portfolios as a strategy to lower English language test anxiety: The case of Chile. *International Journal of Instruction*, 12(1), 181-198. <https://doi.org/10.29333/iji.2019.12112a>
- Coomer, M. N., & Latham-Mintus, K. (2019). Test anxiety: Participation and exclusion beyond the institution. In S. E. Green & D. R. Loseke (Eds.), *New narratives of disability: Constructions, clashes, and controversies*. Emerald Publishing Limited.
- Cooper, P. (2012). *The complete overcoming series: A comprehensive series of self-help guides using cognitive behavioural therapy*. Little, Brown Book Group.
- Craig, S. E., & Stevens, J. E. (2016). *Trauma-sensitive schools: Learning communities transforming children's lives, K-5*. Teachers College Press.
- Crişan, C., & Copaci, I. (2015). The relationship between primary school childrens' test anxiety and academic performance. *Procedia-Social and Behavioral Sciences*, 180, 1584-1589.
- Damer, D. E., & Melendres, L. T. (2011). "Tackling test anxiety": A group for college students. *The Journal for Specialists in Group Work*, 36(3), 163-177.
- Datta, P. (2014). Test anxiety research: Students with vision impairments and students with mild intellectual disabilities. *International Journal of Special Education*, 27(2), 68-74.
- Davidson, M. (2005). Clinical education and evidence-based practice. In M. L. Rose & D. L. Best (Eds.), *Transforming practice through clinical education, professional supervision and mentoring* (pp. 175-184). Elsevier Health Sciences.
- Davies, D. R., Matthews, G., Stammers, R. B., & Westerman, S. J. (2013). *Human performance: Cognition, stress and individual differences*. Taylor & Francis.
- Dawson, P., & Guare, R. (2018). *Executive skills in children and adolescents, third edition: A practical guide to assessment and intervention*. Guilford Publications.
- Dendy, C. Z., Fine, A. H., & Smith, J. (2003). Educational interventions for the adolescent student in the therapy setting. In A. H. Fine & R. A. Kotkin (Eds.), *Therapist's guide to learning and attention disorders* (pp. 175-210). Elsevier Science.

- Deng, L., & Chan, W. (2017). Testing the difference between reliability coefficients alpha and omega. *Educational and psychological measurement*, 77(2), 185-203.
- Department of Education and Skills. (2006). *Circular 0138/2006: Supporting Assessment in Primary Schools*. Dublin 1: Department of Education and Skills
- Department of Education and Skills. (2011). *Circular 0056/2011: Initial Steps in the Implementation of the National Numeracy and Literacy Strategy*. Dublin 1: Department of Education and Skills
- Department of Education and Skills. (2018). *Wellbeing policy statement and framework for practice*. Government of Ireland
- Department of Education and Skills. (2021). *Directory of wellbeing CPD: September 2021-June 2022*. Government of Ireland
- Department of Education and Skills. (2022a). *Circular 0018/2022: Information in relation to standardised testing and other matters- Academic year 2021/2022 and subsequent years*. Dublin 1: Department of Education and Skills
- Department of Education and Skills. (2022b). *Starting a new year: Supporting the wellbeing of school communities: Guidance for schools*. Dublin 1
- Department of Education and Skills, Department of Health, & Health Service Executive. (2015). Well-being in primary schools: Guidelines for mental health promotion. In: Department of Education and Skills.
- Devine, D., Symonds, J., Sloan, S., Cahoon, A., Crean, M., Farrel, E., . . . Hogan, J. (2020). *Children's school lives: An introduction, report no.1* (Children's school lives: National longitudinal cohort study of primary schooling in Ireland, Issue. University College Dublin. <https://a115i2ikt0k3wno971cywnm1-wpengine.netdna-ssl.com/wp-content/uploads/2020/12/CSL-Annual-Report-1-Final.pdf>
- Dooley, B., O' Connor, C., Fitzgerald, A., & A., O. R. (2019). *My world survey 2: The national study of youth mental health in Ireland*. University College Dublin School Of Psychology & t. N. C. f. M. H. Jigsaw. http://www.myworldsurvey.ie/content/docs/My_World_Survey_2.pdf
- Dowker, A. (2019). *Individual differences in arithmetic: Implications for psychology, neuroscience and education*. Taylor & Francis.
- Drake, K. L., Stewart, C. E., Muggeo, M. A., & Ginsburg, G. S. (2015). Enhancing the capacity of school nurses to reduce excessive anxiety in children: Development of the CALM intervention. *Journal of Child and Adolescent Psychiatric Nursing*, 28(3), 121-130. <https://doi.org/10.1111/jcap.12115>

- Dryden, W. (2011). *Manage anxiety through CBT: Teach yourself*. John Murray Press.
- Dugas, M. J., & Robichaud, M. (2007). *Cognitive-behavioural treatment for generalised anxiety disorder: From science to practice*. Routledge.
- Durlak, J. A., Fuhrman, T., & Lampman, C. (1991). Effectiveness of cognitive-behaviour therapy for maladapting children: a meta-analysis. *Psychological bulletin*, 110(2), 204.
- Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development*, 82(1), 405-432.
- Durwin, C. C., & Reese-Weber, M. (2019). *EdPsych modules*. SAGE Publications.
- Edlin, G., & Golanty, E. (2009). *Health and wellness*. Jones & Bartlett Learning.
- Edmunds, M. W., & Mayhew, M. S. (2000). *Pharmacology for the primary care provider*. Mosby.
- English, F. W. (2006). *Encyclopaedia of educational leadership and administration*. SAGE Publications.
- Ergene, T. (2003). Effective interventions on test anxiety reduction: A meta-analysis. *School Psychology International*, 24(3), 313-328.
<https://doi.org/10.1177/01430343030243004>
- Essau, C. A., & Ollendick, T. H. (2012). *The Wiley-Blackwell handbook of the treatment of childhood and adolescent anxiety*. Wiley.
- Eum, K., & Rice, K. G. (2011). Test anxiety, perfectionism, goal orientation, and academic performance. *Anxiety, Stress, & Coping*, 24(2), 167-178.
- Exley, K., Dennick, R., & Fisher, A. (2019). *Small group teaching: Tutorials, seminars and workshops*. Taylor & Francis.
- Eysenck, M. W. (2015). Anxiety and attention. In R. Schwarzer & R. A. Wicklund (Eds.), *Anxiety and self-focused attention* (pp. 125-132). Taylor & Francis.
- Fallon, K., Woods, K., & Rooney, S. (2010). A discussion of the developing role of educational psychologists within children's services. *Educational Psychology in Practice*, 26(1), 1-23. <https://doi.org/10.1080/02667360903522744>
- Fertman, C. I., & Grim, M. (2010). *Health promotion programmes: From theory to practice*. Wiley.
- Fisher, S. (2015). *Stress and strategy*. Taylor & Francis.

- Flippo, R. F., Appatova, V., & Warl, D. M. (2018). Test preparation and test-taking. In R. F. Flippo & T. W. Bean (Eds.), *Handbook of college reading and study strategy research*. Taylor & Francis.
- Frisby, C. L. (2013). *Meeting the psychoeducational needs of minority students: Evidence-based guidelines for school psychologists and other school personnel*. Wiley.
- Fuggle, P., Dunsmuir, S., & Curry, V. (2012). *CBT with children, young people and families*. SAGE Publications.
- Gibbs, G. (1988). *Learning by doing: A guide to teaching and learning methods*. Oxford Further Education Unit.
- Gilmore, C., Göbel, S. M., & Inglis, M. (2018). *An introduction to mathematical cognition*. Taylor & Francis.
- Glazzard, J., & Bancroft, K. (2018). *Meeting the mental health needs of learners 11-18 years*. Critical Publishing.
- Goldin, P. R., Lee, I., Ziv, M., Jazaieri, H., Heimberg, R. G., & Gross, J. J. (2014). Trajectories of change in emotion regulation and social anxiety during cognitive-behavioral therapy for social anxiety disorder. *Behaviour Research and Therapy*, 56, 7-15.
- Gottfried, M. A. (2013). Quantifying the consequences of missing school: Linking school nurses to student absences to standardised achievement. *Teachers College Record*, 115(6), 1-30.
- Gough, D. (2007). Weight of evidence: a framework for the appraisal of the quality and relevance of evidence. *Research Papers in Education*, 22(2), 213-228.
<https://doi.org/10.1080/02671520701296189>
- Government of Ireland. (1999). Social, personal and health education: Primary school curriculum. In: The Stationery Office.
- Grant, C. A., Mericle, B. P., & Royes, S. A. (1994). *Psychiatric nursing*. Delmar Publishers.
- Gregor, A. (2005). Examination anxiety: Live with it, control it or make it work for you? *School Psychology International*, 26(5), 617-635.
- Gullone, E., & Taffe, J. (2012). The Emotion Regulation Questionnaire for Children and Adolescents (ERQ-CA): A psychometric evaluation. *Psychological assessment*, 24(2), 409-417. <https://doi.org/10.1037/a0025777>
- Hagermoser Sanetti, L. M., & Collier-Meek, M. A. (2019). *Supporting successful interventions in schools: Tools to plan, evaluate, and sustain effective implementation*. Guilford Publications.

- Hagermoser Sanetti, L. M., & Kratochwill, T. R. (2009). Toward developing a science of treatment integrity: Introduction to the special series. *School Psychology Review*, 38(4), 445.
- Harrison, J. R., Schultz, B. K., & Evans, S. W. (2017). *School mental health services for adolescents*. Oxford University Press.
- Hartas, D. (2015). Between group differences: Comparing two or more means. In D. Hartas (Ed.), *Educational research and inquiry: Qualitative and quantitative approaches* (pp. 352-377). Bloomsbury Publishing.
- Hasija, S. (1993). *Personality, stress and problem solving*. Northern Book Centre.
- Health Service Executive. (2013). Schools for health in Ireland: Framework for developing a health promoting school. Primary. In: Government of Ireland.
- Heimberg, R. G., & Barlow, G. H. (1995). Children and adolescents: Assessment and treatment. In R. G. Heimberg (Ed.), *Social phobia: Diagnosis, assessment, and treatment* (pp. 387-426). Guilford Publications.
- Hembree, R. (1988). Correlates, causes, effects, and treatment of test anxiety. *Review of Educational Research*, 58(1), 47-77.
- Henn, M., Weinstein, M., & Foard, N. (2005). *A short introduction to social research*. SAGE Publications.
- Hersen, M. (2005). *Encyclopaedia of behaviour modification and cognitive behaviour therapy: Volume II: child clinical applications*. SAGE Publications.
- Heydari, M. G., Shirazi, M., & Sanagouyemoharer, G. R. (2018). The effect of positive psychotherapy in test anxiety among Zahedan students with hemophilia. *Research in Psychotherapy: Psychopathology, Process and Outcome*, 21(1).
- Hobson, S. M. (1996). Test anxiety: Rain or shine! *Elementary School Guidance & Counselling*, 30(4), 316-318.
- Hockenberry, M. J., & Wilson, D. (2014). *Wong's nursing care of infants and children*. Elsevier Health Sciences.
- Hodapp, V., Glanzmann, P. G., & Laux, L. (1995). Theory and measurement of test anxiety as a situation-specific trait. In C. D. Spielberger & P. R. Vagg (Eds.), *Test anxiety: Theory, assessment and treatment* (pp. 47-60). Taylor and Francis.
- Holloway, J. (2019). *Anxiety: Rewire your brain using neuroscience to beat anxiety, fear, worry, shyness, and panic attacks*. Self Publisher.
- Huberty, T. J. (2009). Test and performance anxiety. *Principal Leadership*, 10(1), 12-16.

- Hughes, J. N., & Cavell, T. A. (2013). School-based interventions for aggressive children: Primetime as a case in point. In S. W. Russ & T. H. Ollendick (Eds.), *Handbook of psychotherapies with children and families* (pp. 419-446). Springer US.
- Jackson, D. (2019). *Navigating the maze of research: Enhancing nursing and midwifery practice*. Elsevier Health Sciences.
- Jackson, S. L. (2014). *Research methods: A modular approach*. Cengage Learning.
- Jacob, S., Decker, D. M., & Lugg, E. T. (2016). *Ethics and law for school psychologists*. Wiley.
- Jacobsen, K. H. (2020). *Introduction to health research methods: A practical guide*. Jones & Bartlett Learning, LLC.
- Jamison, J. (2006). *Research methods in psychology for high school students*. iUniverse.
- Jones, L., & Petruzzi, D. C. (1995). Test anxiety: A review of theory and current treatment. *Journal of College Student Psychotherapy, 10*(1), 3-15.
- Joyce-Beaulieu, D., & Sulkowski, M. L. (2015). *Cognitive behavioural therapy in K-12 school settings: A practitioner's toolkit*. Springer Publishing Company.
- Kahneman, D. (1973). *Attention and effort*. Prentice-Hall.
- Kaplan, L. S., & Owings, W. A. (2013). *Culture re-boot: Reinvigorating school culture to improve student outcomes*. SAGE Publications.
- Kaya, A. (2004). Test anxiety and psychopathology in fifth-grade students in Turkey. *International Journal of Educational Reform, 19*-26.
- Kelly, K. (2006). Lived experience and interpretation: The balancing act in qualitative analysis. In M. T. Blanche, M. J. T. Blanche, K. Durrheim, & D. Painter (Eds.), *Research in practice: Applied methods for the social sciences* (pp. 345-369). UCT Press.
- Kendall, J. (2003). Designing a research project: randomised controlled trials and their principles. *Emergency Medicine Journal: EMJ, 20*(2), 164.
- Kennedy, B. (2010). *Anxiety disorders*. Greenhaven Press.
- Kennedy, D. V., & Doepke, K. J. (1999). Multicomponent treatment of a test anxious college student. *Education and Treatment of Children, 20*3-217.
- Khan, A., & Madden, J. (2018). Active learning: a new assessment model that boost confidence and learning while reducing test anxiety. *International Journal of Modern Education and Computer Science, 11*(12), 1.

- King, N. J., Mietz, A., Tinney, L., & Ollendick, T. H. (1995). Psychopathology and cognition in adolescents experiencing severe test anxiety. *Journal of Clinical Child Psychology*, 24(1), 49-54.
- Kivity, Y., Cohen, L., Weiss, M., Elizur, J., & Huppert, J. D. (2021). The role of expressive suppression and cognitive reappraisal in cognitive behavioral therapy for social anxiety disorder: A study of self-report, subjective, and electrocortical measures. *Journal of Affective Disorders*, 279, 334-342.
- Kivunja, C., & Kuyini, A. B. (2017). Understanding and applying research paradigms in educational contexts. *International Journal of Higher Education*, 6(5), 26-41.
- Klenke, K., Martin, S. S., & Wallace, J. R. (2016). *Qualitative research in the study of leadership*. Emerald Group Publishing Limited.
- Kratochwill, T. R. (2003). Task force on evidence-based interventions in school psychology. Retrieved October, 5, 2015.
- Kratochwill, T. R., Cowell, E., Feeney, K., & Hagermoser Sanetti, L. (2003). Behavioural approaches to intervention in educational settings. In R. T. Brown (Ed.), *Handbook of paediatric psychology in school settings* (pp. 521-544). Taylor & Francis.
- Kumar, R. (2010). *Research methodology: A step-by-step guide for beginners*. SAGE Publications.
- Kurth, L., Engelniederhammer, A., Sasse, H., & Papastefanou, G. (2020). Effects of a short mindful-breathing intervention on the psychophysiological stress reactions of German elementary school children. *School Psychology International*, 41(3), 014303432090348-014303432090238. <https://doi.org/10.1177/0143034320903480>
- Lane, D. A., & Corrie, S. (2007). *The modern scientist-practitioner: A guide to practice in psychology*. Taylor & Francis.
- Larsen, K. S. (2017). *Adaptation and well-Being: Meeting the challenges of life*. Taylor & Francis.
- Laurin-Barantke, L., Hoyer, J., Fehm, L., & Knappe, S. (2016). Oral but not written test anxiety is related to social anxiety. *World Journal of Psychiatry*, 6(3), 351.
- Lavell, C., Cadman, J., & Waters, A. (2020). Psychoeducation for exposure therapy. In T. S. Peris, E. A. Storch, & J. F. McGuire (Eds.), *Exposure therapy for children with anxiety and OCD: Clinician's guide to integrated treatment* (pp. 39-68). Elsevier Science.
- Leary, M. R., & Kowalski, R. M. (1997). *Social anxiety*. Guilford Publications.

- Lee, M. T. Y., Wong, B. P., Chow, B. W. Y., & McBride-Chang, C. (2006). Predictors of suicide ideation and depression in Hong Kong adolescents: Perceptions of academic and family climates. *Suicide & Life-Threatening Behaviour, 36*(1), 82-96.
<https://doi.org/10.1521/suli.2006.36.1.82>
- Levine, L. E., & Munsch, J. (2016). *Child development: An active learning approach*. SAGE Publications.
- Lewis-Beck, M., Bryman, E. P. A., Bryman, A. E., Liao, T. F., & Sage Publications, i. (2004). *The SAGE encyclopaedia of social science research methods*. SAGE Publications.
- Liebert, R. M., & Morris, L. W. (1967). Cognitive and emotional components of test anxiety: A distinction and some initial data. *Psychological Reports, 20*(3), 975-978.
- Lobman, C. (2014). "I feel nervous... very nervous" addressing test anxiety in inner city schools through play and performance. *Urban Education, 49*(3), 329-359.
- Loiselle, C. G., Profetto-McGrath, J., Polit, D. F., & Beck, C. T. (2010). *Canadian essentials of nursing research*. Lippincott Williams & Wilkins.
- Lorenzen, T., & Anderson, V. (2018). *Design of experiments: A no-name approach*. CRC Press.
- Lowe, P. A. (2014). The Test Anxiety Measure for Adolescents (TAMA) examination of the reliability and validity of the scores of a new multidimensional measure of test anxiety for middle and high school students. *Journal of Psychoeducational Assessment, 32*(5), 404-416.
- Lowe, P. A. (2019). Expression and level of test anxiety in a sample of elementary students. *Higher Education Studies, 12*(3), 1-9.
- Lowe, P. A., Grumbein, M. J., & Raad, J. M. (2011). Examination of the psychometric properties of the Test Anxiety Scale for Elementary Students (TAS-E) scores. *Journal of Psychoeducational Assessment, 29*(6), 503-514.
- Lowe, P. A., Lee, S. W., Witteborg, K. M., Prichard, K. W., Luhr, M. E., Cullinan, C. M., . . . Janik, M. (2008). The Test Anxiety Inventory for Children and Adolescents (TAICA) examination of the psychometric properties of a new multidimensional measure of test anxiety among elementary and secondary school students. *Journal of Psychoeducational Assessment, 26*(3), 215-230.
- Lyneham, H. J., & Rapee, R. M. (2011). Prevention of child and adolescent anxiety disorders. In W. K. Silverman & A. P. Field (Eds.), *Anxiety disorders in children and adolescents* (pp. 349-366). Cambridge University Press.

- Mandler, G., & Sarason, S. B. (1952). A study of anxiety and learning. *The Journal of Abnormal and Social Psychology*, 47(2), 166.
- Marcz, R. (2017). The mean examiner: How language anxiety affects performance during an exam. In S. L. Krevelj & R. Geld (Eds.), *UZRT 2016: Empirical studies in applied linguistics* (pp. 104-114). Filozofski fakultet u Zagrebu.
- Maria, B. L. (2009). *Current management in child neurology*. BC Decker.
- Marques, L., & Metcalf, E. (2013). *Almost anxious: Is my (or my loved one's) worry or distress a problem?* Hazelden Publishing.
- Mathison, S., & Freeman, M. (2006). Teacher stress and high-stakes testing: How using one measure of academic stress leads to multiple teacher stressors. In R. Lambert & C. McCarthy (Eds.), *Understanding teacher stress in an age of accountability* (pp. 43-64). Information Age Publishing, Inc.
- Mavilidi, M.-F., Hoogerheide, V., & Paas, F. (2014). A quick and easy strategy to reduce test anxiety and enhance test performance. *Applied Cognitive Psychology*, 28(5), 720-726. <https://doi.org/10.1002/acp.3058>
- Mazurek Melnyk, B., & Morrison-Beedy, D. (2018). *Intervention research and evidence-based quality improvement: Designing, conducting, analysing, and funding*. Springer Publishing Company.
- McCambridge, J., Witton, J., & Elbourne, D. R. (2014). Systematic review of the Hawthorne effect: new concepts are needed to study research participation effects. *Journal of clinical epidemiology*, 67(3), 267-277.
- McDonald, A. S. (2001). The prevalence and effects of test anxiety in school children. *Educational Psychology*, 21(1), 89-101.
- McLaughlin, C., & Holliday, C. (2013). *Therapy with children and young people: Integrative counselling in schools and other settings*. SAGE Publications.
- McMillan, J. H. (2017). *Using students' assessment mistakes and learning deficits to enhance motivation and learning*. Taylor & Francis.
- McNabb, D. E. (2020). *Research methods for political science: Quantitative, qualitative and mixed method approaches*. Taylor & Francis.
- Mertens, D. M. (2019). *Research and evaluation in education and psychology: Integrating diversity with quantitative, qualitative, and mixed methods*. SAGE Publications.
- Mills, A. J., Durepos, G., & Wiebe, E. (2009). *Encyclopaedia of case study research*. SAGE Publications.

- Mitchell, B. S., Bruhn, A. L., & Lewis, T. J. (2015). Essential features of tier 2 and 3 school-wide positive behavioural supports. In S. R. Jimerson, M. K. Burns, & A. M. VanDerHeyden (Eds.), *Handbook of response to intervention: The science and practice of multi-tiered systems of support* (pp. 539-562). Springer US.
- Mok, W. S. Y., & Chan, W. W. L. (2016). How do tests and summary writing tasks enhance long-term retention of students with different levels of test anxiety? *Instructional Science*, 44(6), 567-581.
- Morgan, G. A., Gliner, J. A., & Harmon, J. A. (2007). Understanding research methods and statistics: A primer for clinicians. In A. Martin, F. R. Volkmar, & M. Lewis (Eds.), *Lewis's child and adolescent psychiatry: A comprehensive textbook* (pp. 104-123). Wolters Kluwer Health/Lippincott Williams & Wilkins.
- Morrison, M. (2002). What do we mean by educational research? In M. Coleman & A. R. J. Briggs (Eds.), *Research methods in educational leadership and management* (pp. 3-27). SAGE Publications.
- Myers, D., & Dynarski, M. (2003). *Random assignment in programme evaluation and intervention research: Questions and answers*. Institute of Education Sciences, U.S. Department of Education
- Nata, R. V. (2007). *Progress in education*. Nova Science Publishers, Incorporated.
- National Council for Curriculum and Assessment. (2020a). Draft primary curriculum framework: For consultation. In: National Council for Curriculum and Assessment.
- National Council for Curriculum and Assessment. (2020b). Standardised testing 2020/2021. In: National Council for Curriculum and Assessment.
- National Council for Curriculum and Assessment. (n.d.). Primary assessment: Standardised testing. In: National Council for Curriculum and Assessment.
- National Educational Psychological Service. (2007a). Behavioural, Social and Emotional Difficulties: A Continuum of Support: Guidelines for Teachers. In. Dublin: Department of Education and Skills.
- National Educational Psychological Service. (2007b). Special education needs, a continuum of support: Guidelines for teachers. In. Dublin: Department of Education and Skills.
- National Educational Psychological Service. (2020). *Model of service*. Department of Education.
- National Educational Psychological Service. (n.d.). *NEPS Model of Service*. Department of Education and Skills. <https://www.education.ie/en/Schools->

[Colleges/Services/National-Educational-Psychological-Service-NEPS-/Model-of-Service.html](#)

- Naveh-Benjamin, M. (1991). A comparison of training programs intended for different types of test-anxious students: Further support for an information-processing model. *Journal of Educational Psychology*, 83(1), 134.
- Naveh-Benjamin, M., McKeachie, W. J., & Lin, Y.-G. (1987). Two types of test-anxious students: Support for an information processing model. *Journal of Educational Psychology*, 79(2), 131.
- Nilsson, S., Buchholz, M., & Thunberg, G. (2012). Assessing children's anxiety using the modified short state-trait anxiety inventory and talking mats: A pilot study. *Nursing Research and Practice*, 2012.
- O'Donnell, P. S., & Dunlap, L. L. (2019). Teacher acceptability of progressive muscle relaxation in the classroom for the treatment of test anxiety. *Journal of Psychologists and Counsellors in Schools*, 29(2), 151-165.
- O'Leary, M., Lysaght, Z., Nic Craith, D., & Scully, D. (2019). Standardised Testing in English Reading and Mathematics in the Irish Primary School: A Survey of Irish Primary Teachers. In. Dublin: Dublin City University, Centre for Assessment, Research Policy and Practice in Education and Irish National Teachers' Organisation.
- O'Riordan, E. (2019). Over half Leaving Cert pupils have mental or physical health issues due to exam. *The Irish Times*. <https://www.irishtimes.com/news/education/over-half-leaving-cert-pupils-have-mental-or-physical-health-issues-due-to-exam-1.3895005>
- Olatoye, R. A. (2009). Students' test anxiety, motivation for examinations and science achievement in junior secondary schools in Ogun State, Nigeria. *International Journal of Psychology and Counselling*, 1(10), 194-198.
- Olivier, E., Archambault, I., De Clercq, M., & Galand, B. (2019). Student self-efficacy, classroom engagement, and academic achievement: Comparing three theoretical frameworks. *Journal of Youth and Adolescence*, 48(2), 326-340.
- Payne, K. T. (2020). *Preparation for the praxis in speech-language pathology*. Plural Publishing, Incorporated.
- Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review*, 18(4), 315-341.

- Pekrun, R. (2011). Emotions as drivers of learning and cognitive development. In R. A. Calvo & S. K. D'Mello (Eds.), *New perspectives on affect and learning technologies*. Springer New York.
- Pekrun, R. (2016). Academic emotions. In K. R. Wentzel & D. B. Miele (Eds.), *Handbook of motivation at school* (pp. 120-144). Taylor & Francis.
- Pekrun, R. (2017). Achievement emotions. In A. J. Elliot, C. S. Dweck, & D. S. Yeager (Eds.), *Handbook of competence and motivation, second Edition: Theory and application* (pp. 251-271). Guilford Publications.
- Pekrun, R., Cusack, A., Murayama, K., Elliot, A. J., & Thomas, K. (2014). The power of anticipated feedback: Effects on students' achievement goals and achievement emotions. *Learning and Instruction, 29*, 115-124.
<https://doi.org/10.1016/j.learninstruc.2013.09.002>
- Pekrun, R., Goetz, T., Frenzel, A. C., Barchfeld, P., & Perry, R. P. (2011). Measuring emotions in students' learning and performance: The Achievement Emotions Questionnaire (AEQ). *Contemporary Educational Psychology, 36*(1), 36-48.
- Pekrun, R., & Perry, R. P. (2014). Control-value theory of achievement emotions. In *International handbook of emotions in education* (pp. 130-151). Routledge.
- Phares, V. (2003). *Understanding abnormal child psychology*. Wiley.
- Pierangelo, R., & Giuliani, G. (2008). *Teaching students with learning disabilities: A step-by-step guide for educators*. SAGE Publications.
- Pincus, D. B., Korn, A., & DiFonte, M. (2019). Panic disorder. In C. A. Flessner & J. C. Piacentini (Eds.), *Clinical handbook of psychological disorders in children and adolescents: A step-by-step treatment manual* (pp. 204-239). Guilford Publications.
- Pourtaleb, N., Mirnasab, M., & Hadidi, Y. (2018). The Effectiveness of Integrated Training Programme (ITP) in Decreasing Female Students' Test Anxiety: Moderating Effect of Mothers' Perfectionism. *Educational Psychology in Practice, 34*(4), 386-396.
<https://doi.org/10.1080/02667363.2018.1468734>
- Prins, P. J. M., Groot, M. J. M., & Hanewald, G. J. F. P. (1994). Cognition in test-anxious children: The role of on-task and coping cognition reconsidered. *Journal of Consulting and Clinical Psychology, 62*(2), 404.
- Putwain, D., & Daly, A. L. (2014). Test anxiety prevalence and gender differences in a sample of English secondary school students. *Educational Studies, 40*(5), 554-570.
- Putwain, D. W. (2007). Test anxiety in UK schoolchildren: Prevalence and demographic patterns. *British Journal of Educational Psychology, 77*(3), 579-593.

- Putwain, D. W. (2008). Deconstructing test anxiety. *Emotional and Behavioural Difficulties*, 13(2), 141-155.
- Putwain, D. W., & Best, N. (2011). Fear appeals in the primary classroom: Effects on test anxiety and test grade. *Learning and Individual Differences*, 21(5), 580-584.
<https://doi.org/10.1016/j.lindif.2011.07.007>
- Putwain, D. W., & Best, N. (2012). Do highly test anxious students respond differentially to fear appeals made prior to a test? *Research in Education*, 88(1), 1-10.
<https://doi.org/10.7227/RIE.88.1.1>
- Putwain, D. W., & Synes, W. (2020). The four Ws of test anxiety: What is it, why is important, where does it come from and what can be done about it? *Psychologica*, 63(2), 31-51.
- Rana, R. A., & Mahmood, N. (2010). The relationship between test anxiety and academic achievement. *Bulletin of Education and Research*, 32(2).
- Rapee, R., Wignall, A., Spence, S., Lyneham, H., & Cobham, V. (2008). *Helping your anxious child: A step-by-step guide for parents*. New Harbinger Publications.
- Reid, S. (2005). Narrowing the focus: Reflection and choice. In B. Mann (Ed.), *Selected styles in web-based educational research* (pp. 51-69). Information Science Pub.
- Resnik, D. B. (2018). *The ethics of research with human subjects: Protecting people, advancing science, promoting trust*. Springer International Publishing.
- Richardson, F. (2014). Behaviour modification and learning strategies. In H. F. O'Neil (Ed.), *Learning strategies* (pp. 57-78). Elsevier Science.
- Rickels, K., & Rynn, M. A. (2001). What is generalised anxiety disorder? *Journal of Clinical Psychiatry*, 62, 4-14.
- Riddick, B. (2009). *Living With dyslexia: The social and emotional consequences of specific learning difficulties/disabilities*. Taylor & Francis.
- Roberts, M. C., Blossom, J. B., Evans, S. C., Amaro, C. M., & Kanine, R. M. (2018). Advancing the scientific foundation for evidence-based practice in clinical child and adolescent psychology. In M. J. Prinstein (Ed.), *Future work in clinical child and adolescent psychology: A research agenda* (pp. 76-89). Taylor & Francis.
- Robichaud, M., Dugas, M. J., & Antony, M. M. (2015). *The generalised anxiety disorder workbook: A comprehensive CBT guide for coping with uncertainty, worry, and fear*. New Harbinger Publications.
- Rodwell, M. K., & O'Connor, M. K. (1998). *Social work, constructivist research*. Garland Pub.

- Roos, A.-L., Goetz, T., Voracek, M., Krannich, M., Bieg, M., Jarrell, A., & Pekrun, R. (2021). Test anxiety and physiological arousal: a systematic review and meta-analysis. *Educational Psychology Review*, 33(2), 579-618.
- Roselaar, N., Cuadra, M. I., & Lyman, S. (2019). Bias and confounding. In V. Musahl, J. Karlsson, M. T. Hirschmann, O. R. Ayeni, R. G. Marx, J. L. Koh, & N. Nakamura (Eds.), *Basic methods handbook for clinical orthopaedic research: A practical guide and case based research approach* (pp. 23-32). Springer Berlin Heidelberg.
- Rosenthal, G., & Rosenthal, J. A. (2011). *Statistics and data interpretation for social work*. Springer Publishing Company.
- Rosenthal, H. (2013). *Encyclopedia of counselling: Master review and tutorial for the national counsellor examination, state counselling exams, and the counsellor preparation comprehensive examination*. Taylor & Francis.
- Rothman, D. K. (2004). New approach to test anxiety. *Journal of College Student Psychotherapy*, 18(4), 45-60.
- Rumrill, P. D., Cook, B. G., & Stevenson, N. A. (2020). *Research in special education: Designs, methods, and applications*. Charles C Thomas, Publisher, Limited.
- Ruttledge, R., Devitt, E., Greene, G., Mullany, M., Charles, E., Frehill, J., & Moriarty, M. (2016). A randomised controlled trial of the FRIENDS for Life emotional resilience programme delivered by teachers in Irish primary schools. *Educational & Child Psychology*, 33(2), 69-89.
- Salkind, N. J. (2010). *Encyclopaedia of research design*. SAGE Publications.
- Sapp, M. (1991). Hypnotherapy and test anxiety: Two cognitive-behavioural constructs: The effects of hypnosis in reducing test anxiety and improving academic achievement in college students. *Australian Journal of Clinical Hypnotherapy and Hypnosis*.
- Sapp, M. (2010). *Psychodynamic, affective, and behavioural theories to psychotherapy*. Charles C Thomas Publisher, Limited.
- Sapp, M. (2013). *Test anxiety: Applied research, assessment, and treatment interventions*. UPA.
- Sapp, M., & Farrell, W. (1994). Cognitive-behavioural interventions: Applications for academically at-risk and special education students. *Preventing School Failure: Alternative Education for Children and Youth*, 38(2), 19-24.
- Sapp, M., Farrell, W., & Durand, H. (1995). The effects of mathematics, reading, and writing tests in producing worry and emotionality test anxiety with economically and educationally disadvantaged college students. *College Student Journal*.

- Sarason, I. G. (1984). Stress, anxiety, and cognitive interference: reactions to tests. *Journal of Personality and Social Psychology*, 46(4), 929.
- Sarason, I. G., & Palola, E. G. (1960). The relationship of test and general anxiety, difficulty of task, and experimental instructions to performance. *Journal of Experimental Psychology*, 59(3), 185.
- Sarason, S. B., Davidson, K., Lighthall, F., & Waite, R. (1958). A test anxiety scale for children. *Child Development*, 105-113.
- Sarı, S. A., Bilek, G., & Çelik, E. (2018). Test anxiety and self-esteem in senior high school students: a cross-sectional study. *Nordic Journal of Psychiatry*, 72(2), 84-88.
- Savina, E., Fulton, C., & Beaton, C. (2021). *Training teachers in emotional intelligence: A transactional model for elementary education*. Taylor & Francis.
- Sawka-Miller, K. D. (2011). Test anxiety. *Encyclopaedia of Child Behaviour and Development*, 1478-1479.
- Schaefer, A., Matthes, H., Pfitzer, G., & Köhle, K. (2007). Mental health and performance of medical students with high and low test anxiety. *Psychotherapie, Psychosomatik, Medizinische Psychologie*, 57(7), 289-297.
- Schneier, F. R., & Heckelman, L. R. (1995). Diagnostic issues. In R. G. Heimberg (Ed.), *Social phobia: Diagnosis, assessment, and treatment* (pp. 3-20). Guilford Publications.
- Schraw, G. J., & Olafson, L. J. (2007). Assessing teachers' epistemological and ontological worldviews. In M. S. Khine (Ed.), *Knowing, knowledge and beliefs: Epistemological studies across diverse cultures* (pp. 25-44). Springer Netherlands.
- Schuler, P. A. (2000). *Voices of perfectionism: Perfectionistic gifted adolescents in a rural middle school*. National Research Center on the Gifted and Talented.
- Schulz, K. F., & Grimes, D. A. (2002). Sample size slippages in randomised trials: exclusions and the lost and wayward. *The Lancet*, 359(9308), 781-785.
- Schutz, P. A., Distefano, C., Benson, J., & Davis, H. A. (2004). The emotional regulation during test-taking scale. *Anxiety, Stress & Coping*, 17(3), 253-269.
- Segool, N. K., Carlson, J. S., Goforth, A. N., Von Der Embse, N., & Barterian, J. A. (2013). Heightened test anxiety among young children: Elementary school students' anxious responses to high stakes testing. *Psychology in the Schools*, 50(5), 489-499.
- Segool, N. K., Nathaniel, P., Mata, A. D., & Gallant, J. (2014). Cognitive behavioural model of test anxiety in a high-stakes context: An exploratory study. *School Mental Health*, 6(1), 50-61.

- Sharma, S., & Sud, A. (1990). Examination stress and test anxiety: A cross-cultural perspective. *Psychology and Developing Societies*, 2(2), 183-201.
- Sidani, S. (2014). *Health intervention research: Understanding research design and methods*. SAGE Publications.
- Sieber, J. E., O'Neil, H. F., & Tobias, S. (2013). *Anxiety, learning, and instruction*. Taylor & Francis.
- Siegle, D. (2018). Understanding underachievement. In S. I. Pfeiffer (Ed.), *Handbook of giftedness in children: Psychoeducational theory, research, and best practices* (pp. 285-294). Springer International Publishing.
- Silverman, W. K., Kurtines, W. M., Ginsburg, G. S., Weems, C. F., Lumpkin, P. W., & Carmichael, D. H. (1999). Treating anxiety disorders in children with group cognitive-behavioral therapy: A randomised clinical trial. *Journal of Consulting and Clinical Psychology*, 67(6), 995.
- Silvestri, L. A., & Silvestri, A. (2013). *Saunders 2014-2015 strategies for test success: Passing nursing school and the NCLEX exam, 3rd edition*. Elsevier Health Sciences.
- Simos, G., & Hofmann, S. G. (2013). *CBT for anxiety disorders: A practitioner book*. Wiley.
- Smith, E. (2021). *100 ways in 100 days to teach your baby maths: Support all areas of your baby's development by nurturing a love of maths*. Troubador Publishing Limited.
- Smolkowski, K., Strycker, L., & Seeley, J. R. (2013). The role of research in evaluation of interventions for school-related behavioural disorders. In H. M. Walker & F. M. Gresham (Eds.), *Handbook of evidence-based practices for emotional and behavioural disorders: Applications in schools* (pp. 552-566). Guilford Publications.
- Soares, D., & Woods, K. (2020). An international systematic literature review of test anxiety interventions 2011-2018. *Pastoral Care in Education*, 38(4), 311-334.
<https://doi.org/10.1080/02643944.2020.1725909>
- Sousa, D. A. (2016). *How the brain learns*. SAGE Publications.
- Sparfeldt, J. R., Rost, D. H., Baumeister, U. M., & Christ, O. (2013). Test anxiety in written and oral examinations. *Learning and Individual Differences*, 24, 198-203.
- Spence, S. (1994). Preventative strategies. In T. H. Ollendick, N. J. King, & W. Yule (Eds.), *International handbook of phobic and anxiety disorders in children and adolescents*. Springer US.
- Spielberger, C. (1966). Theory and research on anxiety. In C. Spielberger (Ed.), *Anxiety and behaviour* (pp. 3-19). Academic Press.

- Spielberger, C. D. (2010). Test anxiety inventory. *The Corsini Encyclopaedia of Psychology*, 1-1.
- Spielberger, C. D., Anton, W. D., & Bedell, J. (1976). In M. Zuckerman, & CD Spielberger. *The nature and treatment of test anxiety. Emotion and Anxiety: New concepts, methods and applications.*
- Spielberger, C. D., Gonzalez, H. P., Taylor, C. J., Algaze, B., & Anton, W. D. (1978). Examination stress and test anxiety. *Stress and Anxiety*, 5, 167-191.
- Stallard, P. (2012). Cognitive behavioural approaches to mental health difficulties in children. In T. Cole, H. Daniels, & J. Visser (Eds.), *The Routledge international companion to emotional and behavioural difficulties* (pp. 122-130). Taylor & Francis.
- Stallard, P. (2014). *Anxiety: Cognitive behaviour therapy with children and young people.* Taylor & Francis.
- Sterian, M., & Mocanu, M. (2013). Test anxiety. *Euromentor Journal*, 4(3), 75.
- Stoeber, J., Feast, A. R., & Hayward, J. A. (2009). Self-oriented and socially prescribed perfectionism: Differential relationships with intrinsic and extrinsic motivation and test anxiety. *Personality and Individual Differences*, 47(5), 423-428.
- Stroud, K. C. (2013). Methods of assessing learning and study strategies. In D. H. Saklofske, C. R. Reynolds, & V. L. Schwann (Eds.), *The Oxford handbook of child psychological assessment* (pp. 586-613). OUP USA.
- Stuart, K., Maynard, L., & Rouncefield, C. (2015). *Evaluation practice for projects with young people: A guide to creative research.* SAGE Publications.
- Stöber, J., & Pekrun, R. (2004). Advances in test anxiety research. *Anxiety, Stress & Coping: An International Journal*, 17(3), 205-211.
- Sullivan, G. M., & Feinn, R. (2012). Using effect size—or why the P value is not enough. *Journal of Graduate Medical Education*, 4(3), 279-282.
- Sullivan, L. E. (2009). *The SAGE glossary of the social and behavioural sciences.* SAGE Publications.
- Sultan, N. (2018). *Heuristic Inquiry: Researching Human Experience Holistically.* SAGE Publications.
- Talley, N. J., & O'Connor, S. (2010). *Clinical examination: A systematic guide to physical diagnosis.* Elsevier Churchill Livingstone.
- Tate, R., & Perdices, M. (2019). *Single-case experimental designs for clinical research and neurorehabilitation settings: Planning, conduct, analysis and reporting.* Taylor & Francis.

- The Psychological Society Ireland. (2010). *PSI code of professional ethics*.
<https://www.psychologicalsociety.ie/footer/PSI-Code-of-Professional-Ethics-3>
- Thompson, H. R., Duvall, J., Padrez, R., Rosekrans, N., & Madsen, K. A. (2016). The impact of moderate-vigorous intensity physical education class immediately prior to standardised testing on student test-taking behaviours. *Mental Health and Physical Activity, 11*, 7-12.
- Thyer, B. A. (2008). *Preparing research articles*. Oxford University Press, USA.
- Tobias, S. (1985). Test anxiety: Interference, defective skills, and cognitive capacity. *Educational Psychologist, 20*(3), 135-142.
- Troy, A. S., Shallcross, A. J., Brunner, A., Friedman, R., & Jones, M. C. (2018). Cognitive reappraisal and acceptance: Effects on emotion, physiology, and perceived cognitive costs. *Emotion, 18*(1), 58.
- Tryon, G. S. (1980). The measurement and treatment of test anxiety. *Review of Educational Research, 50*(2), 343-372.
- Turner, J. E., & Schallert, D. L. (2001). Expectancy–value relationships of shame reactions and shame resiliency. *Journal of Educational Psychology, 93*(2), 320.
- Ugwuanyi, C. S., Okeke, C. I. O., & Agboeze, M. U. (2020). Management of test anxiety among pupils in basic science using music-based cognitive behaviour therapy intervention: Implication for community development. *Journal of Rational-Emotive & Cognitive-Behaviour Therapy, 1-21*.
- United Nations. (2009). *Convention on the rights of the child* (12, Issue).
<https://www2.ohchr.org/english/bodies/crc/docs/AdvanceVersions/CRC-C-GC-12.pdf>
- University of Cambridge. (n.d.). Rules of thumb on magnitudes of effect sizes. In Cambridge: University of Cambridge.
- Van Blerkom, D. L. (2013). *Orientation to college learning*. Cengage Learning.
- Vargios, E. M. (2007). *Educational psychology research focus*. Nova Science Publishers.
- Von Der Embse, N., Barterian, J., & Segool, N. (2013). Test anxiety interventions for children and adolescents: A systematic review of treatment studies from 2000–2010. *Psychology in the Schools, 50*(1), 57-71.
- von der Embse, N., Jester, D., Roy, D., & Post, J. (2018). Test anxiety effects, predictors, and correlates: A 30-year meta-analytic review. *Journal of Affective Disorders, 227*, 483-493.

- Von der Embse, N. P., Mata, A. D., Segool, N., & Scott, E.-C. (2014). Latent profile analyses of test anxiety: A pilot study. *Journal of Psychoeducational Assessment, 32*(2), 165-172.
- Vye, C., Scholljegerdes, K., & Welch, I. D. (2007). *Under pressure and overwhelmed: Coping with anxiety in college*. Praeger.
- Warner, C. M., Colognori, D., & Lynch, C. (2018). *Helping students overcome social anxiety: Skills for academic and social success (SASS)*. Guilford Publications.
- Warren, S. L., & Sroufe, L. A. (2004). Developmental issues. In T. H. Ollendick & J. S. March (Eds.), *Phobic and anxiety disorders in children and adolescents: A clinician's guide to effective psychosocial and pharmacological interventions*. Oxford University Press.
- Weems, C. F. (2015). Comprehensive school based test anxiety and related problems screening and intervention programme: Group administered test anxiety intervention manual. *New Orleans, LA: University of New Orleans*.
- Weems, C. F., Scott, B. G., Graham, R. A., Banks, D. M., Russell, J. D., Taylor, L. K., . . . Marino, R. C. (2014). Fitting anxious emotion-focused intervention into the ecology of schools: Results from a test anxiety programme evaluation. *Prevention Science, 16*(2), 200-210. <https://doi.org/10.1007/s11121-014-0491-1>
- Weems, C. F., Scott, B. G., Taylor, L. K., Cannon, M. F., Romano, D. M., Perry, A. M., & Triplett, V. (2010). Test anxiety prevention and intervention programmes in schools: Programme development and rationale. *School Mental Health, 2*(2), 62-71. <https://doi.org/10.1007/s12310-010-9032-7>
- Weems, C. F., Taylor, L. K., Costa, N. M., Marks, A. B., Romano, D. M., Verrett, S. L., & Brown, D. M. (2009). Effect of a school-based test anxiety intervention in ethnic minority youth exposed to Hurricane Katrina. *Journal of applied developmental psychology, 30*(3), 218-226. <https://doi.org/10.1016/j.appdev.2008.11.005>
- Weems, C. F., Zakem, A. H., Costa, N. M., Cannon, M. F., & Watts, S. E. (2005). Physiological response and childhood anxiety: Association with symptoms of anxiety disorders and cognitive bias. *Journal of Clinical Child and Adolescent Psychology, 34*(4), 712-723.
- Weis, R. (2020). *Introduction to abnormal child and adolescent psychology*. SAGE Publications.

- Whitaker Sena, J. D., Lowe, P. A., & Lee, S. W. (2007). Significant predictors of test anxiety among students with and without learning disabilities. *Journal of Learning Disabilities, 40*(4), 360-376.
- Whitley, B. E., Kite, M. E., & Adams, H. L. (2013). *Principles of research in behavioural science*. Psychology Press.
- Williams, B. K., & Knight, S. M. (1994). *Healthy for life: Wellness and the art of living*. Brooks/Cole Publishing Company.
- Wine, J. (1971). Test anxiety and direction of attention. *Psychological Bulletin, 76*(2), 92.
- Wine, J. D. (1980). Cognitive-attentional theory of test anxiety. *Test Anxiety: Theory, Research, and Applications, 349-385*.
- World Health Organisation. (2018). *Mental health: strengthening our response*. World Health Organisation. <https://www.who.int/news-room/fact-sheets/detail/mental-health-strengthening-our-response>
- Wren, D. G., & Benson, J. (2004). Measuring test anxiety in children: Scale development and internal construct validation. *Anxiety, stress, and coping, 17*(3), 227-240. <https://doi.org/10.1080/10615800412331292606>
- Yeo, L. S., Goh, V. G., & Liem, G. A. D. (2016). School-based intervention for test anxiety. *Child & Youth Care Forum, 45*(1), 1-17. <https://doi.org/10.1007/s10566-015-9314-1>
- Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Punishment: Issues and Experiments, 27-41*.
- Yew Chye, T. (2008). *Fighting Off Test Anxiety*. Pearson Education South Asia.
- Zamudio, M., Russell, C., Rios, F., & Bridgeman, J. L. (2011). *Critical race theory matters: Education and ideology*. Taylor & Francis.
- Zeidner, & Matthews. (2005). Evaluation anxiety: Current theory and research. In A. J. Elliot & C. S. Dweck (Eds.), *Handbook of competence and motivation, first edition* (pp. 141-166). Guilford Publications.
- Zeidner, M. (2004). Test anxiety. In C. Spielberger (Ed.), *Encyclopaedia of applied psychology* (pp. 545-556). Elsevier Science.
- Zeidner, M. (2006). *Test anxiety: The state of the art*. Springer US.
- Zeidner, M. (2007). Test anxiety in educational contexts: Concepts, findings, and future directions. In *Emotion in education* (pp. 165-184). Elsevier.
- Zeidner, M. (2008). Anxiety revisited: Theory, research, applications. In G. J. Boyle, G. Matthews, & D. H. Saklofske (Eds.), *The SAGE handbook of personality theory and*

assessment: Personality theories and models (volume 1) (pp. 423-446). SAGE Publications.

Zeidner, M. (2014). Test anxiety. In P. Emmelkamp & T. Ehring (Eds.), *The Wiley handbook of anxiety disorders* (pp. 581-596). Wiley.

Zeidner, M., & Matthews, G. (2010). *Anxiety 101*. Springer Publishing Company.

Zengin, M., Yayan, E. H., & Vicnelioğlu, E. (2021). The effects of the COVID-19 pandemic on children's lifestyles and anxiety levels. *Journal of Child and Adolescent Psychiatric Nursing*, 34(3), 236-242.

Zimmerman, A. S., & Kim, J. H. (2021). Excavating and (re)presenting stories: Narrative inquiry as an emergent methodology in the field of adult vocational education and technology. In Information Resources Management Association (Ed.), *Research Anthology on Adult Education and the Development of Lifelong Learners* (pp. 1402-1416). IGI Global.

Zuckerman, M., Anton, W. D., & Bedell, J. (2015). The nature and treatment of test anxiety. In M. Zuckerman & C. D. Spielberger (Eds.), *Emotions and anxiety (PLE: emotion): New concepts, methods, and applications*. Taylor & Francis.

Appendices

Appendix A: List of Excluded Studies

Appendix B: Final List of Studies Included in Review

Appendix C: Summary of Included Studies

Appendix D: Weight of Evidence

Appendix E: Example of Kratochwill's (2003) Coding Protocol Completed for One Study

Appendix F: Invitation Letter to School Principal

Appendix G: Teacher Information Sheet

Appendix H: Teacher Consent Form

Appendix I: Parent Information Sheet

Appendix J: Parent Consent Form

Appendix K: Participant Information Sheet

Appendix L: Participant Assent Form

Appendix M: Ethical Approval Forms

Appendix N: Weems' (2015) Programme Manual

Appendix O: Programme Fidelity Checklist

Appendix P: Children's Test Anxiety Scale

Appendix Q: Adapted Children's Self Efficacy Scale- Academic Achievement

Appendix R: Emotion Regulation Questionnaire for Children and Adolescents

Appendix S: Study Limitations and Considerations for Future Research

Appendix A: List of Excluded Studies*Excluded Studies*

Study	Rationale for Exclusion
(Banks & Smyth, 2015)	No intervention- Criterion 2
(Birturk & Karagun, 2015)	Intervention not school-based- Criterion 2
(Bochis & Sandra, 2018)	No control group- Criterion 7
(Buck, 2016)	No intervention and post- primary participant sample - Criterion 2 and 5 respectively
(Contreras-Soto et al., 2019)	No control group- Criterion 7
(Drake et al., 2015)	Not specific to test anxiety- Criterion 1
(Heydari et al., 2018)	Intervention not school-based- Criterion 2
(Lobman, 2014)	No quantitative measure of TA- Criterion 7
(O'Donnell & Dunlap, 2019)	Qualitative methods used- Criterion 7
(Pekrun et al., 2014)	No measure of TA and post- primary participant sample- Criterion 1 and 5 respectively
(Putwain & Best, 2011)	No intervention- Criterion 2
(Putwain & Best, 2012)	No intervention- Criterion 2
(Ugwuanyi et al., 2020)	Intervention not school-based- Criterion 2

Appendix B: Final List of Studies Included in Review*Included Studies*

References

1. Carsley, D., & Heath, N. L. (2019). Evaluating the effectiveness of a mindfulness colouring activity for test anxiety in children. *The Journal of Educational Research, 112*(2), 143-151.
 2. Carsley, D., Heath, N. L., & Fajnerova, S. (2015). Effectiveness of a classroom mindfulness colouring activity for test anxiety in children. *Journal of Applied School Psychology, 31*(3), 239-255.
 3. Kurth, L., Engelniederhammer, A., Sasse, H., & Papastefanou, G. (2020). Effects of a short mindful-breathing intervention on the psychophysiological stress reactions of German elementary school children. *School Psychology International, 41*(3), 218-238.
 4. Mavilidi, M.-F., Hoogerheide, V., & Paas, F. (2014). A quick and easy strategy to reduce test anxiety and enhance test performance. *Applied Cognitive Psychology, 28*(5), 720-726.
 5. Pourtaieb, N., Mirnasab, M., & Hadidi, Y. (2018). The effectiveness of Integrated Training Programme (ITP) in decreasing female students' test anxiety: Moderating effect of mothers' perfectionism. *Educational Psychology in Practice, 34*(4), 386-396.
 6. Thompson, H. R., Duvall, J., Padrez, R., Rosekrans, N., & Madsen, K. A. (2016). The impact of moderate-vigorous intensity physical education class immediately prior to standardised testing on student test-taking behaviours. *Mental Health and Physical Activity, 11*, 7-12.
 7. Weems, C. F., Scott, B. G., Graham, R. A., Banks, D. M., Russell, J. D., Taylor, L. K., . . . Marino, R. C. (2014). Fitting anxious emotion-focused intervention into the ecology of schools: Results from a test anxiety programme evaluation. *Prevention Science, 16*(2), 200-210. <https://doi.org/10.1007/s11121-014-0491-1>
 8. Yeo, L. S., Goh, V. G., & Liem, G. A. D. (2016). School-based intervention for test anxiety. *Child & Youth Care Forum, 45*(1), 1-17.
-

Appendix C: Summary of Included Studies

No.	Author	Participants	Study design	Intervention	Measures	Outcomes
1.	Carsley & Heath (2019)	N= 152 Education: Elementary Age: (<i>M</i> =10.38) Country: Canada 50% female	Quantitative study Random assignment to control or intervention groups Comparison of pre and post TA within and between groups	Group 1: Mindfulness art activity (mandala) Group 2: Free drawing colouring (control group)	State Trait Anxiety Inventory form (STAI-C)	Significant decrease in TA from pre to post interventions No significant gender differences
2.	Carsley et. al. (2015)	N= 52 Education: Private Elementary Age: (<i>M</i> =10.92) Country: Canada 53.8% female	Quantitative study Comparison of TA pre and post interventions within and between groups Randomly assigned to groups	Group 1: Structured mandala condition Group 2: Free drawing/ colouring (control group)	State-Trait Anxiety Inventory for Children- State form (STAIC-S)	Significant decrease in TA following mandala condition in both males and females Males benefited from both conditions however females benefited only from mandala condition in reducing TA
3.	Kurth et al. (2020)	N= 106 Education: Elementary School Age: 5-11 (<i>M</i> = 8.7) Country: Germany	Quantitative study Each class randomly divided into 2-3 groups and randomly assigned to the intervention or control group	Group 1: Intervention group-with breathing exercise Group 2: Comparison group-without breathing exercise	Smartband to measure physiological responses to stress induced by the testing situation	Intervention group showed significant difference in stress arousal compared to control group. Treatment effect remained significant despite stress level acting as a moderator

		58% female	Stress arousal peaks monitored and compared between groups			
4.	Mavilidi et al. (2014)	N= 117 Education: Primary Age: 11-12 (M= 11.59) Country: Greece 56% female	Quantitative Study Participants divided into low, medium or high level of anxiety groups and randomly assigned to intervention or control groups Between-group comparison of TA	Group 1: Experimental condition- looking ahead strategy Group 2: Control condition- non-looking ahead strategy	Cognitive Anxiety Test Questionnaire	Main effect of test strategy was not significant
5.	Pourtaleb et al. (2018)	N= 60 Education: Primary Age: Not reported Country: Iran 100% female	Quantitative study Participants randomly assigned to two experimental and two control groups TA measured and compared between and within-groups	Experimental Group 1: 15 participants with perfectionist mothers received intervention (integrated training programme) Experimental Group 2: 15 participants without perfectionist mothers received intervention (integrated training programme)	Spielberger Test Anxiety Inventory	Significant reduction from pre to post intervention Intervention group displayed significantly lower TA levels than control group at post-intervention

				Control Group 1: 15 participants perfectionist mothers- no intervention		
				Control Group 2: 15 participants without perfectionist mothers – no intervention		
6.	Thompson et al. (2016)	N= 791 Education: Public elementary Age: M= 10.6-10.7 Country: United States of America 47% female	Quantitative study Cluster randomised control trial TA measured post-intervention and compared between groups Standardised test scores compared between and within-groups	Group 1: P.E. intervention condition Group 2: No intervention control condition	Children’s Test Anxiety Scale (CTAS) Teacher survey of student test-taking behaviour, including 8 questions from the CTAS Standardised test results for reading and math	No statistically significant differences in TA between intervention and control group No statistically significant difference between groups in the change in test scores following intervention
7.	Weems et al. (2014)	N = 259 (Treatment groups) N = 163 (No treatment follow-up group)	Quantitative study Random assignment to groups where possible	Group 1: Group-based cognitive-behaviour therapy (CBT) intervention Group 2: Waitlist control group	Test Anxiety Scale for Children	Statistically significant reduction in TA in both groups. Group 1 demonstrated steeper decrease in TA levels with a significant effect of treatment group. Further but

		<p>Education: Public schools-grades 3-12 Age: 8-17 Country: USA 62% Female</p>	<p>TA compared within and between groups with five timepoints: pre-intervention, post-intervention and three follow-ups.</p>	<p>No treatment group: A non-test-anxious group compared for follow-up measures</p>		<p>smaller declines in TA observed at follow-ups.</p>
8.	Yeo et al. (2016)	<p>N= 115 Education: Elementary Age: 9-12 (M= 10.15) Country: Singapore 39% female</p>	<p>Quantitative Study Quasi-experimental - non-randomised assignment of participants to intervention and control groups</p> <p>TA levels compared at baseline, post and follow-up and compared between-groups</p>	<p>Group 1: Classroom-based CBT intervention Group 2: No intervention control group</p>	<p>Children's Test Anxiety Scale</p>	<p>Significant reduction in TA levels from baseline to follow-up but no significant differences from pre to post intervention</p>

Appendix D: Weight of Evidence**Weight of Evidence A- Methodological Quality**

Kratochwill's (2003) Coding Protocol for group-designs was used to evaluate WoE A. This protocol addresses general characteristics and involves eight key features used to determine the quality of evidence of intervention studies; (II-A) Measurements, (II- B) Comparison Group, (II- C) Primary/Secondary Outcomes, (II- D) Educational/Clinical Significance, (II-E) Identifiable components, (II-F) Implementation Fidelity, (II-G) Replication, (II-H) Site of Implementation, (II-I) Follow-up Assessment. The protocol was amended for the suitability of this review; excluded sections are outlined in the table below.

Excluded sections from Kratochwill's (2003) Coding Protocol.

Excluded sections	Rationale
I: General Characteristics	Components within these sections are outlined in the summary of studies (Appendix C) and in the review
II- C: Primary/ Secondary Outcomes	Outcomes are not evaluated at this stage of weighting of evidence (methodological quality only)- evaluated at later stage
II- D: Clinical Significance	Evaluated in the review
II- E: Identifiable Interventions Components	As stated, outcomes and effects evaluated and discussed later
II- F: Implementation Fidelity	There were no issues of implementation fidelity as in all included studies the interventions were either delivered by the researchers, provided a training package or required no training
II- G: Replication Studies	No included studies were replication studies

The included sections used of Kratochwill's (2003) coding protocol to evaluate WoE are outlined in the tables below.

Measurements

Weighting	Criteria
High (3)- Strong evidence	<ul style="list-style-type: none"> - Reliability of scores for primary outcome measures should be $\geq .85$ - Reliability of measure in the current population under study must be reported - Data collected using multiple methods - Data collected from multiple sources - Validity of measures are reported
Medium (2)- Promising evidence	<ul style="list-style-type: none"> - Reliability of scores for primary outcome measures should be $\geq .70$ - Reliability of measure in the current population under study must be reported - Data collected using either multiple methods or multiple sources - Case for validity doesn't need to be presented - Above criteria must be met for 75% of primary outcome measures when multiple primary outcomes
Low (1)- Weak evidence	<ul style="list-style-type: none"> - Reliability of scores for primary outcome measures should be $\geq .50$ - Reliability of measure in the current population under study must be reported - Data may be collected using either multiple methods or multiple sources but not required - Case for validity doesn't need to be presented - Above criteria must be met for 50% of primary outcome measures when multiple primary outcomes

Comparison Groups

Weighting	Criteria
High (3)- Strong evidence	<ul style="list-style-type: none"> - At least one active comparison group - Group equivalence through random assignment of participants to intervention groups - Evidence for change agents being counterbalanced - Equivalent Mortality and low attrition at post
Medium (2)- Promising evidence	<ul style="list-style-type: none"> - At least a no intervention group for comparison - Evidence for two of the following: <ul style="list-style-type: none"> - Group equivalence through random assignment of participants to intervention groups - Evidence for change agents being counterbalanced - Equivalent Mortality and low attrition at post
Low (1)- Weak evidence	<ul style="list-style-type: none"> - A comparison group and at least one of the following: <ul style="list-style-type: none"> - Group equivalence through random assignment of participants to intervention groups - Evidence for change agents being counterbalanced - Equivalent Mortality and low attrition at post

Site of Implementation

Weighting	Criteria
High (3)- Strong evidence	<ul style="list-style-type: none"> -Study must be carried out in a public or alternative school setting - Home school partnerships also receive this rating if the intervention is initiated by the school in an outreach effort if it is a public or alternative school
Medium (2)- Promising evidence	<ul style="list-style-type: none"> -Study must be conducted in private or charter school setting -Studies will also receive this rating if the type of school is unknown

- Home school partnerships also receive this rating if the intervention is initiated by the school in an outreach effort if it is a private or charter school

Low (1)- Weak evidence -The intervention does not have to implemented in a school setting but could be implemented with minor adjustments

Follow-up Assessment

Weighting	Criteria
High (3)- Strong Evidence	<ul style="list-style-type: none"> -Follow-up assessments must have been conducted at multiple time intervals e.g. six months, one year - Follow-up must be carried out with all participants from the original sample - Similar measures must be used to analyse primary/secondary outcomes
Medium (2)- Promising Evidence	<ul style="list-style-type: none"> - Follow-up must be carried out at least once e.g. 6 months - Follow-up must include the majority of participants from the original sample - Similar measures must be used to analyse primary/secondary outcomes
Low (1)- Weak Evidence	<ul style="list-style-type: none"> - Follow-up must be carried out at least once e.g. 6 months - Follow-up must include some participants from the original study

WoE A Scores

Study	Measurements	Comparison Groups	Site of Implementation	Follow-up assessment	Overall WoE A (average of scores)
Carsley & Heath (2019)	1 (low)	2 (medium)	3 (high)	0 (no evidence)	1.5 (medium)
Carsley & Heath (2015)	1 (low)	2 (medium)	2 (medium)	0 (no evidence)	1.25 (low)
Kurth et al. (2020)	0 (no evidence)	2 (medium)	2 (medium)	0 (no evidence)	1 (low)
Mavilidi et al. (2014)	0 (no evidence)	2 (medium)	2 (medium)	0 (no evidence)	1 (low)
Pourtaleb et al. (2018)	1 (low)	2 (medium)	2 (medium)	0 (no evidence)	1.25 (low)
Thompson et al., (2016)	0 (no evidence)	2 (medium)	3 (high)	0 (no evidence)	1.25 (low)
Weems et al. (2014)	1 (low)	2 (medium)	3 (high)	1 (low)	1.75 (medium)
Yeo et al. (2016)	1 (low)	2 (medium)	3 (high)	2 (medium)	2 (medium)

Note: Low = < 1.4, Medium = 1.5-2.4, High = > 2.5

Weight of Evidence B (WoE B)- Methodological Relevance

To determine WoE B scores for the methodological relevance of studies in the effectiveness of school-based interventions in reducing TA, Gough's (2007) Weight of Evidence Framework was utilised. Gough states that WoE B involves making a judgment on a study in relation to the appropriateness of the research design in answering the review question.

Weighting and Criteria for WoE B

Weighting	Criteria
High (3)- Strong evidence	<ul style="list-style-type: none"> - Must include an 'active' control group for comparison - Participants must be randomly assigned to intervention or control groups -Must measure and report primary outcomes pre and post intervention for both intervention and control groups -Must use appropriate statistical analysis to measure primary outcomes
Medium (2)- Promising evidence	<ul style="list-style-type: none"> -Must include a control group for comparison -Must include random assignment of participants -Appropriate pre and post measures used to determine primary outcomes
Low (1)- Weak evidence	<ul style="list-style-type: none"> -Must include control condition -Random assignment of participants is not required -Primary outcome measured pre and post intervention

Rationale for Weight of Evidence B

- The use of a control group is required to determine how much of the effect is due to the intervention rather than alternative factors such as researcher effects. An alternative intervention group, as opposed to a non-intervention group, would provide the strongest evidence as it determines if the intervention is superior to an alternative intervention (Balakrishnan, 2014).

- Participants must be randomly assigned to group conditions to ensure equality between groups and reduce the likelihood of group differences influencing the results for the effectiveness of the intervention.
- As the review is specifically examining the effects of the intervention using quantitative methods, the review is relying on the use of appropriate statistical analysis to determine these effects. Differences in TA must be measured pre and post intervention to determine the effectiveness of the intervention in reducing TA.

WoE B Scores for Included Studies

Study	WoE B Score
Carsley & Heath, 2019	3 (high)
Carsley & Heath, 2015	3 (high)
Kurth et al., 2020	0 (no evidence)
Mavilidi et al., 2014	0 (no evidence)
Pourtaleb et al., 2018	2 (medium)
Thompson et al., 2016	0 (no evidence)
Weems et al., 2014	2 (medium)
Yeo et al., 2016	1 (low)

Note: Low = < 1.4, Medium = 1.5-2.4, High = >2.5

Weight of Evidence C (WoE C)- Relevance of Evidence

To evaluate WoE C, Gough's (2007) framework was used. This framework describes WoE C as a specific judgment of the focus of the evidence in relation to answering the review question.

Weighting	Criteria
High (3)- Strong evidence	<ul style="list-style-type: none"> - TA must be measured pre and post intervention as the primary outcome - A test must be immediately followed by the intervention - Must also include an alternative intervention control group for comparison - Participant sample is collected from various schools

- Medium (2)- Promising evidence
- TA must be measured pre and post intervention as one of the primary outcomes
 - Participants are currently studying in a school where they are subject to tests
 - Must include a control group for comparison, can be either alternative intervention or non-intervention
 - Participant sample is collected from various classes in schools
- Low (1)- Weak evidence
- TA must be measured as an outcome
 - Must include control group for comparison
 - Participant sample is collected from a school where they are subject to tests during their time there

Note: Low = < 1.4, Medium = 1.5-2.4, High = >2.5

Rationale for weight of evidence C

- As the research question is focused on reducing TA, TA should be measured as an outcome, preferably a primary outcome.
- As TA is a type of anxiety which is induced specifically by a fear of being tested, participants must be subject to test situations to determine the effect of the intervention of this test-specific type of anxiety. It is preferable that the intervention takes place immediately prior to a test, when TA is likely to occur naturally, to determine the effectiveness of the intervention.
- Participants should be gathered from a variety of settings, which are either from various schools or various classes within a school to ensure the generalisability of the results.
- Must include an alternative intervention, or at the very least a non-intervention group, to determine how much of the effect is due to the intervention (Balakrishnan, 2014).

Weight of Evidence C Scores for Included Studies

Study	WoE C Score
Carsley & Heath, 2019	3 (high)
Carsley & Heath, 2015	2 (medium)
Kurth et al., 2020	1 (low)

Mavilidi et al., 2014	1 (low)
Pourtaleb et al., 2018	2 (medium)
Thompson et al., 2016	1 (low)
Weems et al., 2014	2 (medium)
Yeo et al., 2016	2 (medium)

Appendix E: Example of Kratochwill's (2003) Coding Protocol Completed for One Study

As mentioned, amendments were made to the coding protocol for this review; this example is based on the adapted version for an evaluation of methodological quality (WoE A).

Coding Protocol: Group-Based Design

Domain: School-wide and classroom-based programs

Name of Coder(s): [REDACTED]

Full Study Reference in APA format: Carsley, D., Heath, N. L., & Fajnerova, S. (2015). Effectiveness of a classroom mindfulness colouring activity for test anxiety in children. *Journal of Applied School Psychology, 31*(3), 239-255.

Intervention Name (description from study): Mindfulness-based structured colouring activity

Study ID Number (Unique Identifier): _____

Type of Publication: Journal Article

II. Key Features for Coding Studies and Rating Level of Evidence/ Support**(3=Strong Evidence 2=Promising Evidence 1=Weak Evidence 0=No Evidence)****A. Measurement** (answer A1 through A4)

A1. Use of outcome measures that produce reliable scores for the majority of primary outcomes. The table for Primary/Secondary Outcomes Statistically Significant allows for listing separate outcomes and will facilitate decision making regarding measurement (select one of the following)

- A1.1 Yes
 A1.2 No
 A1.3 Unknown/unable to code

A2. Multi-method (select one of the following)

- A2.1 Yes
 A2.2 No
 A2.3 N/A
 A2.4 Unknown/unable to code

A3. Multi-source (select one of the following)

- A3.1 Yes
 A3.2 No
 A3.3 N/A
 A3.4 Unknown/unable to code

A4. Validity of measures reported (select one of the following)

- A5.1 Yes validated with specific target group
 A5.2 In part, validated for general population only
 A5.3 No
 A5.4 Unknown/unable to code

Rating for Measurement (select 0, 1, 2, or 3): 3 2 1 0

B. Comparison Group**B1. Type of Comparison Group** (select one of the following)

- B1.1 Typical contact
 B1.2 Typical contact (other) specify:
 B1.3 Attention placebo
 B1.4 Intervention elements placebo
 B1.5 Alternative intervention
 B1.6 Pharmacotherapy B1.1
 B1.7 No intervention
 B1.8 Wait list/delayed intervention
 B1.9 Minimal contact
 B1.10 Unable to identify comparison group

Rating for Comparison Group (select 0, 1, 2, or 3): 3 2 1 0

B2. Overall confidence rating in judgment of type of comparison group (select one of the following)

- B2.1 Very low (little basis)
 B2.2 Low (guess)
 B2.3 Moderate (weak inference)
 B2.4 High (strong inference)
 B2.5 Very high (explicitly stated)
 B2.6 Unknown/Unable to code

B3. Counterbalancing of Change Agents (answer B3.1 to B3.3)

- B3.1 By change agent
 B3.2 Statistical
 B3.3 Other

B4. Group Equivalence Established (select one of the following)

- B4.1 Random assignment
 B4.2 Posthoc matched set
 B4.3 Statistical matching
 B4.4 Post hoc test for group equivalence

B5. Equivalent Mortality (answer B5.1 through B5.3)

- B5.1 Low Attrition (less than 20% for Post)
 B5.2 Low Attrition (less than 30% for follow-up)
 B5.3 Intent to intervene analysis carried out
 Findings _____

H. Site of Implementation**H1. School (if school is the site, select one of the following options)**

- H1.1 Public
 H1.2 Private
 H1.3 Charter
 H1.4 University Affiliated
 H1.5 Alternative
 H1.6 Not specified/unknown

H2. Non School Site (if it is a non school site, select one of the following options)

- H2.1 Home
 H2.2 University Clinic
 H2.3 Summer Program
 H2.4 Outpatient Hospital
 H2.5 Partial inpatient/day Intervention Program
 H2.6 Inpatient Hospital
 H2.7 Private Practice
 H2.8 Mental Health Center
 H2.9 Residential Treatment Facility
 H2.10 Other (specify): _____
 H2.11 Unknown/insufficient information provided

Rating for Site of Implementation (select 0, 1, 2, or 3): 3 2 1 0

I. Follow Up Assessment

Timing of follow up assessment: specify N/A

Number of participants included in the follow up assessment: specify N/A

Consistency of assessment method used: specify N/A

Rating for Follow Up Assessment (select 0, 1, 2, or 3): 3 2 1 0

Appendix F: Invitation Letter to School Principal

Dear Principal/Chairperson,

My name is Alison Mc Fadden, and I am a Trainee Educational Psychologist. I am currently in my second year of the Professional Doctorate in Child and Educational Psychology in Mary Immaculate College, Limerick. I have received Garda Vetting from Mary Immaculate College and have completed Child First training. I have experience working with children in my role as an SNA prior to this course and on professional placement with an Early Intervention and Disability Team and I am currently on placement with the National Educational Psychological Service (NEPS). As part of the requirements of this course, I am conducting a research study on the effectiveness of a school-based test anxiety programme at primary-level. Test anxiety refers to the extreme fear of assessment situations which can result in reduced academic performance. Symptoms can include increased heart rate and muscle tension, attention, concentration, and memory difficulties, worry and feelings of hopelessness and inadequacy, as well as task avoidance and fidgeting. In a survey of Irish primary teachers in 2019, three out of four teachers reported pupils experiencing extreme anxiety in the lead up to standardised tests which suggests that test anxiety is an increasing issue at primary-level in Ireland. Therefore, it is my aim to evaluate the effectiveness of a school-based intervention which involves educating children on test anxiety and teaching emotion management skills to cope with and prevent test anxiety. I have completed a risk assessment and safeguarding statement in line with Mary Immaculate College's Safeguarding Children Policy and Procedures to ensure protection of participants. This study has been approved by the Mary Immaculate Research Ethics Committee.

I would be grateful if you would consider permitting this study to be carried out in your school and to allow me to seek the consent of pupils and parents to participate in this study. Additionally, I will require consent from the school's Board of Management to proceed with this study.

The intervention will take place over two 6-week periods which will involve two groups receiving the intervention in 6 sessions each in a classroom setting; each session will last approximately one hour. I hope to run the intervention with two 4th class groups with one group (Group 1) receiving the intervention for the first 6 weeks while the other class (Group 2) take part in the intervention in a second 6-week block. This will require random assignment of each class to either Group 1 or Group 2. As part of this study pupils in both groups will be asked to complete three self-report measures of test anxiety, academic self-efficacy and emotion regulation at 3 time-points; prior to the intervention, at the end of the intervention and at a 2-month follow-up, as well as a demographic questionnaire at the beginning. Parents of pupils will also be asked to complete a parent questionnaire based on their perception of their child's test anxiety at the same 3 time-points.

I hope to begin this study in September 2021. The starting point of this will be to complete a pilot study, which is a small study prior to a larger study to determine the reliability of two measures, the Children's Test Anxiety Scale and a parent measure of their child's levels of test anxiety. This will involve administering two self-report measures in a classroom setting with a similar age group to the participant sample (i.e. with another 4th class if possible or a 5th class sample) and giving participants the parent measure to take home for parents to complete. This will then be followed by delivering the intervention in 6 weekly sessions to Group 1 (the intervention group) in September/October 2021. Considering time for school holidays and the researcher's work/university schedule, it is planned that the intervention will end in November 2021. A 2-month follow-up will then be carried out in January 2022 which will involve participants and their parents filling out measures again. Once the 2-month follow-up has been conducted, Group 2 (waitlist control group) will then receive the intervention which I hope to carry out in January 2021.

Participation in this study is entirely voluntary and participants may withdraw from the study at any time without consequence. If a child does not give consent to participate, does not have parental consent, or chooses to withdraw from the study at any stage, they may still participate in the programme activities, however, they will not be required to complete questionnaires or self-report measures. Alternatively, they may choose to do another activity in the classroom during sessions, such as colouring/drawing or reading a book. All data gathered will be confidential, identifying information such as names, dates of birth or address will not be requested on questionnaires. Instead, participants will be assigned a number which will appear on their questionnaires. Participant numbers will be stored beside their names in a password-protected excel file which only the researcher will have access to. Participants will have the right to withhold self-report measures and questionnaires if they change their mind during the study. All data gathered in the study will be kept confidential in a locked folder and on a password-protected laptop. In accordance with Mary Immaculate College's retention policy, all data will be stored indefinitely. Once I have gathered the data in this study, I will analyse and interpret the findings to be reported in my final thesis; this may be disseminated by way of research articles or other formats.

In relation to confidentiality, all information discussed in sessions will be kept confidential with certain exceptions where confidentiality may have to be broken i.e. if there is a risk of harm to a child or other persons, or if information relating to a crime is disclosed. Confidentiality will be broken if this happens whereby the researcher will disclose this information to a designated liaison person (Carrie Ryan, Geraldine Brosnan or Paula Hourigan) in Mary Immaculate College and in the school, if required. Additionally, children may express more generalised anxiety during discussion of test anxiety; if this happens, the protocol will be to ensure that any child experiencing generalised anxiety is reported to the child's teacher to ensure that they receive the appropriate supports and that parents are informed.

This will be communicated to and agreed with participants in the information sheet and consent form prior to commencement of the intervention.

If you have any further queries or questions in relation to this study, please do not hesitate to contact me at the following email: 19183135@micstudent.mic.ul.ie. Alternatively, you can contact my supervisors:

Dr. Fionnuala Tynan, email: Fionnuala.Tynan@mic.ul.ie

Dr. John Perry, email: John.Perry@mic.ul.ie

If you have any concerns about this study and wish to contact someone independent, you may contact:

MIREC Administrator

Research and Graduate School

Mary Immaculate College

South Circular Road

Limerick

Telephone: 061 204980

Email: mirec@mic.ul.ie

Thank you for taking the time to read this letter. If you are interested in allowing me to run this intervention in your school, I would be grateful if you contact me by email on 19183135@micstudent.mic.ul.ie. Once I receive your email, I would be happy to discuss this further in person or over the phone.

Yours Sincerely,

Alison Mc Fadden

(Trainee Educational Psychologist)

Contact number: 0868432496

Email: 19183135@micstudent.mic.ul.ie

Appendix G: Teacher Information Sheet



Test Anxiety and The Effectiveness of School-Based Intervention at Primary-level: Information Sheet for Teachers

My name is Alison Mc Fadden, and I am a Trainee Educational Psychologist. I am currently in my third year of the Professional Doctorate in Child and Educational Psychology in Mary Immaculate College, Limerick. As part of the requirements of this course, I am conducting a research study on the effectiveness of a school-based test anxiety programme at primary-level. Test anxiety refers to the extreme fear of assessment situations which can result in reduced academic performance. Symptoms can include increased heart rate and muscle tension, difficulties with attention, concentration, and memory, worry and feelings of hopelessness and inadequacy, as well as task avoidance and fidgeting.

What does the study involve?

The study will involve the implementation of a classroom-based programme which aims to reduce test anxiety in six weekly sessions, each session will last approximately 45 minutes to one hour. The sessions will include education on anxiety in general and test anxiety, followed by the teaching of emotion regulation skills to reduce test anxiety; emotion regulation refers to the ability to effectively manage and respond to emotions. Participants will also be asked to complete three self-report scales of test anxiety, emotion regulation, and self-efficacy at 3 time-points: the beginning of the programme, the end of the programme and 2-months following the programme, as well as a demographic questionnaire at the beginning. Additionally, parents will be asked to complete a parent questionnaire to examine their perception of the level of their child's test anxiety at the three time points. Two groups will receive the intervention; Group 1 will receive the intervention first and will be required to complete measures and questionnaires at 3-time points, at the beginning of the intervention, at the end of the intervention and at a 3-month follow-up while Group 2 will also be asked to complete these at the same time, however, will not receive the intervention until 3-month follow up has been conducted in September 2021.

What is your role in this study?

As part of this study, I hope to ensure equal access and participation of all consenting participants to the study. In doing so, I hope to consult with you, as the participants' teacher, to discuss any children who may require additional support to engage with the intervention, including reading and filling out questionnaires.

Furthermore, children may express more generalised anxiety during discussion of test anxiety; if this happens, the protocol will be to report this to you, as the child's teacher. This protocol aims to ensure that any child experiencing generalised anxiety receives the appropriate supports in school under the Wellbeing Policy Framework and Statement for Practice (2018), and so that you can report to parents so that they are informed of this.

Protection of the children in your class

I have completed a risk assessment and safeguarding statement prior in line with Mary Immaculate College's Safeguarding Children Policy and Procedures to this study. All information discussed in sessions with your class will be kept confidential with certain exceptions where confidentiality may have to be broken i.e. if there is a risk of harm to a child or other persons, or if information relating to a crime is disclosed. Confidentiality will be broken if this happens whereby the researcher will disclose this information to the designated liaison person (Carrie Ryan, Geraldine Brosnan or Paula Hourigan) in Mary Immaculate College. If a child in your class does not give consent to participate in this study, does not have parental consent, or chooses to withdraw from the study at any stage, they may still participate in the programme activities, however, they will not be required to complete questionnaires or self-report measures. Alternatively they may choose to do another activity in the classroom during sessions, such as colouring/drawing or reading a book.

Consent

If you agree to facilitate this study, please sign the consent form attached.

Further queries/questions

If you have any further queries or questions in relation to this study, please do not hesitate to contact me (Alison Mc Fadden) at the following email: 19183135@micstudent.mic.ul.ie. Alternatively, you can contact my supervisors:

Dr. Fionnuala Tynan, email: Fionnuala.Tynan@mic.ul.ie

Dr. John Perry, email: John.Perry@mic.ul.ie

If you have any concerns about this study and wish to contact someone independent, you may contact:

MIREC Administrator

Research and Graduate School

Mary Immaculate College

South Circular Road

Limerick

Telephone: 061 204980

Email: mirec@mic.ul.ie

Appendix H: Teacher Consent Form



Consent Form- Teacher

<p>I agree to the allocation of time for this intervention to be run in my classroom once a week for 6 weeks for approximately 45 minutes to an hour per session.</p>	
<p>I agree to inform the researcher if any participants have needs that would require additional support for equal participation in the study.</p>	
<p>I agree to receive information from the researcher in relation to any child experiencing generalised anxiety and to act accordingly to ensure that their needs are supported and that parents are informed.</p>	

Signed: _____

Print name: _____

Date: _____

Appendix I: Parent Information Sheet



Test Anxiety and The Effectiveness of School-Based Intervention at Primary-level: Information Sheet for Parents

My name is Alison Mc Fadden, and I am a Trainee Educational Psychologist (TEP). As part of the requirements for the Doctorate in Child and Educational Psychology in Mary Immaculate College, I am carrying out a study on the effectiveness of a school-based intervention for test anxiety at primary-level. I would like to invite the participation of your child in this study.

What will the study involve?

The study will involve the implementation of a classroom-based programme which aims to reduce test anxiety in six weekly sessions, each session will last approximately 45 minutes to one hour and will be delivered by the researcher (Alison Mc Fadden, Trainee Educational Psychologist). Test anxiety refers to the extreme fear of assessment situations which can result in reduced academic performance. The sessions will include education on anxiety in general and test anxiety, followed by the teaching of emotion regulation skills to reduce test anxiety; emotion regulation refers to the ability to effectively manage and respond to emotions. Your child will be asked to complete a questionnaire on demographic information such as age, gender, and country of origin at the beginning of the study. They will also be asked to complete three self-report scales of test anxiety, emotion regulation, and self-efficacy (belief in own abilities) at 3 time-points: the beginning of the programme, the end of the programme and 2-months following the programme. Additionally, you will be asked to complete a parent questionnaire to examine your perception of your child's test anxiety at the three time points.

As part of the delivery of this programme, your child will be randomly assigned to one of two groups. If they are in Group 1, they will begin the programme in September/October 2021. If they are in Group 2, they will be asked to complete questionnaires and scales at the same time as Group 1, however, they will not receive the programme until January 2022.

Voluntary participation

Your child has been asked to take part in this study as the target population for this study is children currently attending 4th class at primary-level. Participation in this study is completely voluntary, you do not have to agree to the participation of your child in this study. Additionally, you or your child can withdraw from participation in the study at any time without consequence. If you or your child do not give consent to participate, or your child chooses to withdraw from the study at any stage, they may still participate in the programme activities, however, they will *not* be required to complete questionnaires or self-report measures. Alternatively, they may choose to do another activity in the classroom during sessions, such as colouring/drawing or reading a book.

Confidentiality

All data gathered in this study will remain confidential, identifying information such as names or date of births will *not* be requested on questionnaires, this includes the name of the school. Instead, each participant will be assigned a number which will appear on their questionnaire; participants numbers will be stored beside each name on a password-protected document by the researcher. This is for the purpose of identifying corresponding questionnaire scores at the three timepoints, and corresponding parent and child questionnaires for analysis. The information gathered during the study will be reported anonymously in the thesis written by the researcher.

All information discussed in sessions with your child will be kept confidential with certain exceptions where confidentiality may have to be broken i.e. if there is a risk of harm to your child or other persons, or if information relating to a crime is disclosed. Confidentiality will be broken if this happens whereby the researcher will disclose this information to a designated liaison person (Carrie Ryan, Geraldine Brosnan or Paula Hourigan) in Mary Immaculate College and in the school, if required. Additionally, children may express more generalised anxiety during discussion of test anxiety; if this happens, the protocol will be to ensure that any child experiencing generalised anxiety is reported to their class teacher to ensure they receive the appropriate supports, as many teachers are now equipped to deal with general anxiety under the Wellbeing Policy Framework and Statement for Practice (2018). Your child's teacher will be required to inform you, as your child's parent/guardian, if this happens.

Storage of data

All hardcopies of gathered information will be stored in a locked folder where only the researcher will have access to it. Data will then be transferred to an online excel file and to statistical software for storage and analysis, this data will be stored on a password-protected laptop which only the researcher will have access to. The results from the data collected will be written up in a thesis as part of the course requirements discussed above, this may be disseminated by way of research articles or other formats. In accordance with Mary Immaculate College's data retention policy, data will then be stored indefinitely.

Access to data

Your child has the right to withdraw from the study at any time and withhold data (i.e. questionnaires and scales) before the final analysis of data.

Are there risks in taking part?

As the topic of test anxiety is a sensitive topic, this may cause upset or distress for either yourself or your child. If you or child decide that you are no longer happy to take part in the study, you can withdraw at any time. You can also find support from the following helplines and websites below:

- Parentline is a national, confidential helpline which offers support to parents in all aspects of being a parent including around anxiety and exam stress.

Link to website: <https://www.parentline.ie/>

Contact number: 1890 927 277

- You can also find resources and useful links for guidance and support from the government's Parent and Family Support publication at the following link:

<https://www.gov.ie/en/publication/5825b6-parenting-and-family-support/>

- Childline is also a useful resource for any child experiencing distress or anxiety. They can free phone, text or live message from the website using the following details:

Free phone: 1800 66 66 66

Free text: 50101

Live message: <https://www.childline.ie/login/>

What are the benefits of taking part?

Test anxiety has become an increasing issue at Primary-level in Ireland; in 2019, three out of four teachers reported in a survey that pupils experience extreme anxiety in the lead up to standardised tests. It has also been found that test anxiety can result in reduced performance and therefore, test results can provide an inaccurate depiction of a child's learning and progress. The aim of this programme is to reduce and prevent test anxiety in the classroom and subsequently, improve academic performance of those experiencing test anxiety. Although not all children may experience test anxiety, this programme may still be beneficial in preventing potential future test anxiety by equipping children with coping strategies for the increased demands of examinations in the future.

Equal participation

To ensure equal access to the programme and that all participants can participate equally, you will be asked to indicate if your child has any additional needs or diagnoses that would require additional support (e.g. with reading or filling in forms, questionnaires etc.) to enable full participation. This information will also be asked of your child's teacher to ensure appropriate support throughout the programme.

Who has approved this study?

This study has been approved by the Mary Immaculate Research Ethics Committee and by the Principal of Scoil Íosagáin, Sinead McLaughlin. A risk assessment and safeguarding statement in line with Mary Immaculate College's Safeguarding Children Policy and Procedures have also been completed to protect participants in this study.

Further queries/questions

If you have any further queries or questions in relation to this study, please do not hesitate to contact me (Alison Mc Fadden) at the following email: 19183135@micstudent.mic.ul.ie. Alternatively, you can contact my supervisors:

Dr. Fionnuala Tynan, email: Fionnuala.Tynan@mic.ul.ie

Dr. John Perry, email: John.Perry@mic.ul.ie

If you have any concerns about this study and wish to contact someone independent, you may contact:

MIREC Administrator

Research and Graduate School

Mary Immaculate College

South Circular Road

Limerick

Telephone: 061 204980

Email: mirec@mic.ul.ie

Consent

If you agree to the participation of your child in this study, please sign the attached consent form and return to the school.

Appendix J: Parent Consent Form**Parent Consent Form**

Dear Parent/Guardian,

Your child has been invited to participate in a study on the effectiveness of a school-based programme for test anxiety. Please read the attached information sheet carefully before signing this consent form and tick the relevant box below to indicate your understanding and consent.

I have read and understand the information sheet attached to this form.	
I understand the purpose of the study and what the programme will involve.	
I understand that I have the right to refuse the participation of my child in this study. My child has the right to withdraw from this study at any time.	
I understand that I will be required to fill out a questionnaire based on my child's level of test anxiety at 3 timepoints.	
I understand that identifying information (e.g. name, address) will be not appear on questionnaires but will be stored on a password-protected document.	
I understand that my child has the right to refuse submission of data (i.e. questionnaire and self-report measures).	
I have read and understood the risks and benefits associated with participating in this study.	
I understand that my child's teacher will disclose if my child has any additional needs or diagnoses that will require additional support to engage with the study (e.g. with reading or filling in forms, questionnaires etc.).	
I understand that all information will be kept confidential, however, this may have to be broken if there is risk to my child or other persons, if they disclose details of a crime or if they report more generalised anxiety.	
I understand that the information gathered in this study will be written up in a thesis and data will be stored indefinitely.	

Child's name (in print): _____

Parent Signature: _____

Name in print: _____

Date: _____

If you have agreed to the participation of your child in this study, please indicate below if your child has any additional needs or diagnoses that may influence their ability to participate in this study and the supports required to ensure their equal participation:

Appendix K: Participant Information Sheet



Test Anxiety and The Effectiveness of School-Based Intervention at Primary-level: Information Sheet for Participants

My name is Alison Mc Fadden, I am carrying out a project on test anxiety and how we can reduce this in the classroom. Test anxiety is the feeling of extreme fear of tests. I am doing this project as part of my training to become an Educational Psychologist. This means that I am training to work in helping children with lots of things, such as their worries or things they may find difficult in school.

What will the project involve?

This project will involve 6 sessions, once a week, which will take 45 minutes to 1 hour and will be carried out in your classroom. Your class will be chosen for either Group 1 or Group 2. If you are in Group 1, you will take part in the project first. You will be asked to answer some questions in the first and last sessions about how tests make you feel and how you deal with your feelings. If you are in Group 2, you will be second to receive these sessions, however, you will be asked to answer the same questions as Group 1 at the same time as they do.

What will I learn?

You will learn about what test anxiety means and ways to deal with this, such as breathing exercises and imagining facing your fears. Even if you do not feel you have experienced test anxiety, you can still learn about how to cope with exam fears and worries for the future.

Do I have to take part?

Participation in this study is voluntary, this means that you do not have to agree to take part in this study if you do not want to. If you do not want to take part, you may still join in with the programme activities, however, you will not have to complete the questionnaires. You may also choose to do another activity in the classroom during sessions, such as colouring/drawing or reading a book.

What if I do not like the programme or want to change my mind?

As our fears and worries can be a hard thing to talk about, it is possible that you may feel worried or upset during the study. If you change your mind during the project, you can stop taking part at any time by telling the researcher (Alison). You do not have to hand up your questionnaires. Also, if you feel you have been upset by anything in this study, you can seek help by telling an adult, including the researcher. You can also contact Childline which is a service who helps children using the details on the next page:

Free phone: 1800 66 66 66

Free text: 50101

Live message: <https://www.childline.ie/login/>

If you have any concerns about this study and wish to contact someone independent, you may contact:

MIREC Administrator

Research and Graduate School

Mary Immaculate College

South Circular Road

Limerick

Telephone: 061 204980

Email: mirec@mic.ul.ie

Will my information be kept private?

When filling out your questionnaires, you will not have to put your name on them, instead you will be given a unique number to put on the questionnaire. Everything we discuss in sessions will be kept private, however, if there is a chance of harm to you or other persons, or if information about a crime is given then this will be passed onto another adult. Also, if it appears that you are experiencing anxiety, which means extreme fears and worries, around things other than tests and exams, this will be discussed with your teacher who will also tell your parent/guardian. This is to make sure that you are protected and that you receive help if needed.

What will happen this information?

This information will be used by the researcher (Alison) to write up a project for college, your name or the name of your school will not be written in this report.

Want to take part?

If you would like to take part in this project, please fill in the form on the next page.

Appendix L: Participant Assent Form



Please read the following sentences and tick each box if you agree.

	<p>I have read and understand the information about this project in the sheet provided.</p>	
	<p>I agree to take part in this project.</p>	
	<p>I understand that I do not have to take part in this study if I do not want to and that I can stop taking part in this study at any time, including during the project.</p>	
	<p>I understand and agree that all information I provide will remain private, however, I know that if there is a risk of harm to myself or others or if I discuss a crime, this agreement may be broken. I also understand that if I discuss more general worries, that this will be discussed with my teacher who will also tell my parent/guardian.</p>	
	<p>I understand that information from the project will be written up by the researcher (Alison).</p>	

Name _____

Today's date _____

Appendix M: Ethical Approval Forms



**Mary Immaculate College
Research Ethics Committee**
MIREC-4: MIREC Chair Decision Form

APPLICATION NO.

A20-067 FINAL

1. PROJECT TITLE

Test Anxiety and The Effectiveness of School-Based Intervention at Primary Level

2. APPLICANT

Name:	Alison McFadden
Department / Centre / Other:	EPISE
Position:	Postgraduate Researcher

3. DECISION OF MIREC CHAIR

<input type="checkbox"/>	Ethical clearance through MIREC is required.
<input type="checkbox"/>	Ethical clearance through MIREC is not required and therefore the researcher need take no further action in this regard.
<input checked="" type="checkbox"/>	Ethical clearance is required and granted. Referral to MIREC is not necessary.
<input type="checkbox"/>	Ethical clearance is required but the full MIREC process is not. Ethical clearance is therefore granted if required for external funding applications and the researcher need take no further action in this regard.
<input type="checkbox"/>	Insufficient information provided by applicant / Amendments required.

4. REASON(S) FOR DECISION

A20-067 – Alison Mc Fadden - Test Anxiety and The Effectiveness of School-Based Intervention at Primary Level

I have reviewed this application and I believe it satisfies MIREC requirements. It is, therefore, approved.

5. DECLARATION (MIREC CHAIR)

Name (Print):	Dr Marie Griffin
Signature:	
Date:	12 th January 2021



**Mary Immaculate College
Research Ethics Committee**
MIREC-4: MIREC Chair Decision Form

APPLICATION NO.

A20-067 1 st Amendment Request
--

1. PROJECT TITLE

Test Anxiety and The Effectiveness of School-Based Intervention at Primary Level
--

2. APPLICANT

Name:	Alison McFadden
Department / Centre / Other:	EPISE
Position:	Postgraduate Researcher

3. DECISION OF MIREC CHAIR

<input type="checkbox"/>	Ethical clearance through MIREC is required.
<input type="checkbox"/>	Ethical clearance through MIREC is not required and therefore the researcher need take no further action in this regard.
<input checked="" type="checkbox"/>	Ethical clearance is required and granted. Referral to MIREC is not necessary.
<input type="checkbox"/>	Ethical clearance is required but the full MIREC process is not. Ethical clearance is therefore granted if required for external funding applications and the researcher need take no further action in this regard.
<input type="checkbox"/>	Insufficient information provided by applicant / Amendments required.

4. REASON(S) FOR DECISION

A20-067 Amendment 1 – Alison Mc Fadden - Test Anxiety and The Effectiveness of School-Based Intervention at Primary Level
--

I have reviewed this request and I believe it satisfies MIREC requirements. It is, therefore, approved.

5. DECLARATION (MIREC CHAIR)

Name (Print):	Dr Marie Griffin
Signature:	
Date:	17 th August, 2021

Appendix N: Weems' (2015) Programme Manual

Group Administered Test Anxiety Intervention Manual

Carl Weems, Ph.D.

Created while I was at the University of New Orleans

Correspondence To:

Carl F. Weems, Department of Human and Family Studies, 4380 Palmer, Iowa State University, Ames, IA

50011. E-mail: cweems@iastate.edu

Overview

This manual describes an intervention designed to reduce school-aged students' (~10 -17 years) test anxiety. The content is based on the empirical literature on test anxiety and effective treatments for childhood anxiety problems (Weems et al., 2010). The intervention has been administered as a selective targeted intervention and was designed for implementation with students in group and classroom guidance settings who have been identified as potentially benefiting from test anxiety reduction skills (can also be administered individually with minor modification). The protocol was designed to teach basic emotion management skills to help students who may do poorly on regular or standardized tests such as state wide testing because of test anxiety and the content was developed from Cheek et al. (2002) and Silverman et al. (1999). It is designed for flexible use and integration with other programs. Students identified as potentially benefiting from a test anxiety intervention can be administered the protocol in as few as four sessions and with content typically elaborated and reiterated to extend the number of sessions to 5-6. The intervention may also be administered as a universal programme (i.e., offered to all).

Theory: Research suggests the effectiveness of exposure and relaxation training (cognitive and behavioural therapy techniques) for school aged students in the reduction of anxiety disorder symptoms broadly (e.g., Silverman et al. 1999). Evidence also suggest that incorporating reinforcing activities (such as art and/or music techniques) with the anxiety reducing strategies may provide additional support and an element of fun that can help motivate students to use the strategies (Cheek et al., 2002). Cognitive modification strategies (e.g., negative thought restructuring techniques) are not the focus in this manual because research suggests that test anxious students may be further hampered by techniques that could inadvertently promote additional off task thoughts (King et al., 1995; Prins et al., 1994). The goal of the test anxiety intervention is facilitating an optimal level of arousal (i.e., not over aroused/motivated and not under aroused/motivated) paired with on task thoughts. However, cognitive strategies such as praise from the therapist and other strategies that promote the students' self-efficacy are used. The students' parents can also be given information and materials to use at home. This manual content has empirical support in the form of a randomized waitlist controlled treatment study among high risk 9th graders (Weems et al., 2009). Additional effectiveness data on youth in grades 3-12 suggest the protocol may be effective in reducing test anxiety in earlier grades (Weems et al., 2014). Weems et al., 2014 screened at-risk youth (N = 1,048) from urban public schools and 325 with elevated test

anxiety were offered the intervention in one of two waves (immediate intervention versus waitlist). The intervention was associated with decreases in test anxiety, anxiety disorder, and depression symptoms and improvement in a sense of control for older youth. Results suggest high participant satisfaction and growth curve analysis of follow-up assessments (end of the year, the next school year, and a subsequent school year) demonstrated positive developmental trajectories consistent with predictions (e.g., initial change in test anxiety predicted change in other symptoms) suggesting maintenance of gains and possible long term prevention benefits.

Important to note: This manual is not intended to serve as a comprehensive guide to effective intervention with youth, it assumes familiarity and experience with CBT techniques for anxiety reduction in youth or supervision of manual implementation by someone who does. The techniques for delivering session content should also be tailored to the developmental level of the groups. For example, more drawing and picture-based content for younger groups and more verbal discussion and text content for older groups. The techniques for delivering content across the age range in this study have been previously validated with similar effects across age groups (see Silverman et al., 1999).

References:

- Cheek, J. R. Bradley, L. J., Reynolds, J., & Coy, D. (2002). An intervention for helping elementary students reduce test anxiety. *Professional School Counseling, 6*, 162-164.
- King, N. J., Mietz, A., Ollendick, T. H., & Tinney, L. (1995). Psychopathology and cognition in adolescents experiencing severe test anxiety. *Journal of Clinical Child Psychology, 24*, 49-54.
- Prins, P. J., Groot, M. J., & Hanewald G. J. (1994). Cognition in test-anxious children: the role of on-task and coping cognition reconsidered. *Journal of Consulting and Clinical Psychology, 62*, 404-409.
- Silverman, W. K., Kurtines, W. M., Ginsburg, G. S., Weems, C. F., Lumpkin, P. W., & Carmichael, D. H. (1999). Treating anxiety disorders in children with group cognitive-behavioural therapy: A randomized clinical trial. *Journal of Consulting and Clinical Psychology, 67*, 995-1003.
- Weems, C. F., Taylor, L. K., Costa, N. M., Marks, A. B., Romano, D. M., Verrett, S. L., & Brown, D. M. (2009). Effect of a school-based test anxiety intervention in ethnic minority youth exposed to Hurricane Katrina. *Journal of Applied Developmental Psychology, 30*, 218-226.
- Weems, C. F., Scott, B. G., Graham, R. A., Banks, D. M. Russell, J.D. Taylor, L. K., Cannon, M. Varela, R. E., Scheeringa, M. S. Perry A. M. & Marino, R., C. (2014). Fitting Anxious Emotion Focused Intervention into the Ecology of Schools: Results from a Test Anxiety Programme Evaluation. *Prevention Science*. DOI: 10.1007/s11121-014-0491-1
- Weems, C. F., Scott, B. G., Taylor, L. K., Cannon, M. F., Romano, D. M., Perry A. M., & Triplett, V. (2010). Test anxiety prevention and intervention programmes in schools: Programme development and rationale. *School Mental Health, 2*, 62-71

Procedure

Screen (i.e., pre test) for test anxiety (e.g., use the TASC- see end of manual for suggested measures).

Students who have elevated test anxiety or otherwise identified need (e.g., who do not meet passing rates or who exhibited or reported extreme feelings of anxiety and stress) are then administered the following programme. Re-administration of the test anxiety scales at the immediate beginning and end of the programme will help evaluate effectiveness and identify areas of potential improvement in the programme. The programme can also be given to all students regardless of test anxiety levels but may be less relevant for those who are not anxious.

First session:

Overview: in session 1, participants learn about anxiety in general, test anxiety, and evaluation fears. The counsellor tries to foster a sense of universality about evaluation fears and participants develop a test anxiety hierarchy. The overall rationale is explained as being able to take tests with an optimal level of motivation. The counsellor tries to develop the sense of universality regarding test anxiety by explaining that the students should not feel bad about being scared of tests because many people are scared of tests. The counsellor also presents a cognitive and behavioural conceptualization of anxiety. The counsellor explains the importance of emotion regulation and facing the fear. Students work on their test anxiety exposure hierarchy. This hierarchy consists of various aspects of testing (e.g., from the announcement of the test, to studying, to turning in the test). Discussion of the students' reactions to the items serves as a focal point for the group relaxation techniques. The counsellor generates a consensus in the group about the scariest parts on the test hierarchy in an effort to facilitate group practices.

See integrity checklist for each specific component (end of document).

Administer the pre-treatment assessment

Develop Rapport and the Common Goals—Administer Pre-treatment Assessment.

It is very important for students to develop a sense of universality about test anxiety and for the therapist to develop rapport with the group. Use your training to develop rapport: Empathy, Openness, Congruence.

Orientation to the programme. Give a brief overview of the programme and create the sense that this programme is a joint effort between the counsellor and the children to help them do well. Ask if they have any questions? Convey the following: It is important to ask questions. Any time they are unclear, they should ask. Don't be shy or embarrassed. No question is stupid. The counsellor/therapist wants to help them. That is what the therapist is here for. Provide an overall rationale as to why this programme is being suggested and begin to discuss the goal: I.e., being able take tests with an optimal level of motivation. Develop a sense of universality regarding test anxiety—that they should not feel bad about being scared of tests many people are: explain that the skills they learn in this programme will help them feel less nervous.

Present a conceptualization of test anxiety. It is explained to the child that when we are afraid or anxious, fear or anxiety usually shows itself in three ways: (1) bodily reactions, such as heart beating faster than usual, stomach aches, sweat, etc., (2) talking to oneself, ("I might

fail"), and (3) actions or behaviours, i.e., we avoid the feared object or event. Anxiety is natural (it helps motivate us); however, when it gets too intense it can interfere with our ability to test well.

Explain the importance of exposure or approach behavior. Generally, it is explained that when individuals engage in avoidance of tests it is maintained because the individuals never learn that "there is really nothing to be afraid or anxious of." The best way to learn that "there is really nothing to be afraid or anxious of," is to gradually expose oneself, i.e., to approach the fearful object/event or to perform the anxious behavior. (The analogy of "getting back on a bicycle after falling off" should be made.)

Again---Make sure students don't feel they are different or silly for being nervous when taking tests. Lots of kids are. Develop sports and entertainment analogies –query about the groups interests in this regard.

Students are then administered a test anxiety exposure hierarchy based on the work of Kennedy and Doepke (1999) . This hierarchy consists of 13 items. The students are asked to respond to each item by individually recording the level of his or her anxiety on a scale of 0 (representing no anxiety) to 8 (representing debilitating anxiety). Discussion of the students' reactions to the items serves as a focal point for the group relaxation techniques. Generating a general consensus in the group about the scariest parts on the test hierarchy will facilitate group practices, but individual differences should not be neglected.

A copy of the hierarchy is attached in the appendices of this manual.

It is very important to complete the hierarchy in the first session and develop a sense of universality and for the therapist to develop rapport with the group. Completion of relaxation exercises training can be started in the second session—if necessary.

Introduce the children to systematic relaxation techniques. Using the tag-along method, first demonstrate a few of the relaxation exercises, then practice with the child/children, then have the child do the exercises on his/her own. Use of sport, music or acting analogies is useful. Breathing in Karate, basketball. Control of muscles in gymnastics or dance, etc.

Relaxation Exercises

Why is doing this important? Because if our bodies feel tensed up or weird, then sometimes it makes taking tests harder to do. If we learn how to relax our bodies, it can help us do our best.

First--Muscle Relaxation:

1. Get comfortable sitting!
2. Make each part of your body tense by squeezing all the muscles tight (not too tight), hold it for three seconds, and then relax. Start with your face, then do:
 - neck and shoulders
 - arms and hands

- chest and stomach
- legs and back
- feet
- finally, all body parts together.

Second--Breathing Exercises:

1. Now that you know how to relax your muscles, practice relaxing your “insides”.
2. Pretend you are a balloon that gets big when it fills up with air and shrinks when the air goes out.
3. Put one hand on your stomach and one hand on your chest.
4. Breath in through your nose, blowing up like a balloon, hold it for three seconds (or otherwise comfortable period of time), and let the air out--pushing out all of the nervous and worried feelings in your body.
5. Repeat one time. This is “relaxing”.
6. As you feel more and more relaxed, imagine yourself in your favourite place--a place where you feel completely relaxed and happy. As you feel your body get looser, and you are thinking about your favourite place, you will notice that both sides of your mouth begin to go up and a big smile will show up on your face.

Remember: Just as you can make your body tense, you can make it relax by practicing these exercises.

Finish session by explaining that we will begin to use these skills to help us deal with test anxiety. Explain that whenever you feel anxious or tense you can use these skills. Tell students that they can practice these for next time. Use examples from individuals in the media the group is familiar with (eg., a sport or music celebrity and their use of emotion control)

Therapist should complete treatment integrity checklist for session 1—any items not completed should be completed in the next session.

Second session:

Overview: In session 2, participants learn or consolidate learning relaxation techniques depending on progress in session 1, begin self-efficacy building, and discuss general test-taking skills (e.g., learning when to study --don't put off or “avoid” studying, the importance of approach and non-avoidance is discussed). The relaxation techniques are first practiced during imagined exposure to the initial or lower level hierarchy items. The goal of these sessions is to create a context where students can practice their relaxation skills while “facing” or approaching items on the stimulus hierarchy. The counsellor reviews why facing your fears is important and other content from session 1. Children are reinforced through verbal praise on any progress in learning by the counsellor and by their peers in the group.

See integrity checklist for each specific component.

Review or begin relaxation exercises.

Demonstrate the tag-along method again especially if the children need more help learning these exercises. Ask the children if they found them easy to do. Were they useful to feel relaxed? If not, offer to modify the exercises. For instance, a particular child may benefit from a less complex set of relaxation techniques. In any case, these exercises are “successful” when the child is able to wind down and relax.

As a mnemonic device, students can be taught to "Stop, Drop, and Roll." This technique utilizes the well-known fire safety precautions that many children are taught in schools. Specifically, the students can be instructed that when they physically feel the "fire" of anxiety and stress, they can "stop" (actually put down their pencils and place their hands on the table while concentrating on the coolness of the surface i.e., get comfortable). Then they "drop" their heads forward, and "roll" them around gently while taking three deep breaths. They can also use other parts of the relaxation techniques at this time.

Sneaky Muscles: Some students may wish to use covert methods of doing muscle relaxation and breathing exercises. These can be described as “sneaky muscles”. Students should be instructed in ways to do progressive muscle relaxation and breathing that are covert.

Mr Universe: Some students may think the relaxation exercises are silly or “stupid. These beliefs should be directly addressed. One way is to describe the “cool” aspects of muscle and breathing control (again via individuals the group may look up to). For example, body builders use similar techniques to strengthen and define muscles. Martial arts experts use similar techniques in Karate, Judo, and Tai Kwan Do.

Discussion of the hierarchy. The relaxation techniques are practiced during imagined exposure to the initial or lower level hierarchy items. As the counsellor reads each item, the group members imagine the item and can say "Fire!" when they started to feel anxiety, and then the group practices the "Stop, Drop, and Roll/Relaxation techniques.

If progress on relaxation is apparent in vivo exposure (e.g., a mock test) combined with practicing the relaxation techniques can be used. The goal of these session is to create a context where students can practice their relaxation skills while “facing” approaching items on the stimulus hierarchy. Explain to the group that facing your fears is important. As part of a group discussion ask them “why is facing your fears important” Lead the discussion to the point that facing your fears head on helps you to learn the skills to deal with hard problems in life like important tests.

Reinforce through verbal praise any progress in learning --focus on progress and not perfection in learning this material.

Therapist should complete treatment integrity checklist for session 2—any items not completed should be completed in the next session.

Third session:

Overview: Session 3 focuses on test anxiety hierarchy exposure tasks (i.e., gradual exposure to items on the hierarchy combined with relaxation such as “Sitting in a classroom waiting for the test to begin, and they hand you the test.”) with more self-efficacy building, and more relaxation training practice. If progress is apparent along the hierarchy, actual exposure (e.g.,

a mock test) combined with practicing the relaxation techniques is used in this session. The counsellor introduces the concept of self-evaluation as deciding whether or not one is pleased or satisfied with his/her own work. Sports and entertainment examples are used to help youth learn the concept of shaping (or gradual learning).

See integrity checklist for each specific component.

Review previous information. Review in older groups may be in the form of verbal discussion. Review in younger groups may take the form of handouts depicting information or drawing exercises.

Review relaxation exercises.

Demonstrate the relaxation tag-along methods again especially if the children need more help learning these exercises. Ask the children if they found them easy to do. Were they useful to feel relaxed? If not, offer to modify the exercises. For instance, a particular child may benefit from a less complex set of relaxation techniques. In any case, these exercises are “successful” when the child is able to wind down and relax.

Students continue to practice the relaxation techniques during imagined exposure to the rest of the hierarchy items. Counsellor/therapist’s should use verbal praise to reinforce progress and help build self-efficacy. If progress is apparent in vivo exposure (e.g., a mock test) combined with practicing the relaxation techniques can be used in this session.

Introduce the idea of self-evaluation by describing self-evaluation as deciding whether or not you are pleased or satisfied with your own work. Suggest that people actually evaluate themselves and reward or punish themselves for their own behavior, although they may not even know it. Provide concrete examples of both situations, drawn from the children’s own experiences (e.g., for a Little Leaguer "Imagine that you hit a home run that helped your team win the game. "Were you successful at what you tried to do? What would you be thinking about? What would you do afterwards? How would you feel?") Be sure to point out that it would not be reasonable to expect to hit a home run every time you are at bat. This is not justification to punish your-self.

Help children make a list of possible rewards that the child might do or give to him/herself (emphasize verbal self-rewards). Examples might include: "Great Job!" "I really handled that well," "I can handle it if I try," "Good going," "I'm really proud of myself," Explain the importance in believing in yourself—because you can always do your best. Utilize sports and entertainment analogies developed in earlier sessions based on groups interests.

Explain how we should reward or praise ourselves for even partial successes. Specifically, the counselor reminds the children how it was just explained that one might praise oneself for hitting home run but also pointed out that we should not necessarily expect ourselves to hit a home run every time up at bat. Indeed, if we had never hit a baseball before, we would be just as happy the first time at bat to simply hit the ball. If we did this the first time, we ought to praise our self. And if next time at bat we hit a single, we would praise our self for this, and then next time a double... until home run. That is, we will praise ourselves for partial successes as well, not just for the home run.

Practice with making self-evaluations and rewarding oneself. Have each child describe an experience that is at least somewhat successful and have them demonstrate how they would evaluate and reward their performance.

Emphasize to the child that they should praise themselves for any progress and not just the times when they do everything perfectly. Focus on the idea that no one does everything perfectly and not doing something 100% perfectly should not mean that you should not praise yourself.

As a reward and way to help foster this, the students can use any type of art materials to create self-portraits depicting themselves as calm and successful during the test. For older youth they can draw a picture of what they think success for them will mean (a nice car, home, etc). Have older youth write a short story (paragraph) about being successful at school this year. If a student wishes let them share their success story. These drawings along with a copy of the relaxation techniques can be sent home with the students so that their parents can assist them with their relaxation.

Therapist should complete treatment integrity checklist for session 3—any items not completed should be completed in the next session.

Fourth Session:

Overview: Session 4 reviews content from sessions 1 to 3 and focuses on conducting more exposure tasks similar to the above at higher points of the hierarchy. There is also additional self-evaluation and self-efficacy training. Any information from sessions 1-3 not yet presented should be at this time.

See integrity checklist for each specific component.

Review the previous information on relaxation and self-praise.

This session should be used to finalize teaching the relaxation and self-praise material and as an extension of the third session where the counsellor uses in vivo exposure (e.g., a mock test) combined with practicing the relaxation techniques.

For example, students can practice teach their classmates the "Stop, Drop, and Roll" techniques (monitored and **assisted by the counsellor**).

An additional reinforcing project similar to the drawing at the end of session 3 can be implemented at the end of session 4.

Therapist should complete treatment integrity checklist for session 4—any items not completed should be completed in the next session.

Fifth Session:

Overview: Session 5 includes any final exposures that are needed on the hierarchy, and there is a graduation party/meeting where participants are praised for completing the programme. This session is for the completion of any exposures and to reinforce learning on all the topics. Be sure to review all content and have final discussion about progress and skills learned. All group members should have completed between 4 and 5 exposures and be able to use relaxation exercises.

See integrity checklist for each specific component.

1. Discussion and review of all major components, brain-storming with group about continuing the progress made.

2. Say goodbye, thank them for their participation and reinforce the learning by the children with verbal praise for all their progress.

This session is used to finish any components that have not been completed, to review, and to provide a reward/graduation session such as handing out certificates of completion, etc.

3. Administer the post intervention assessment. Depending on the activity planned it is best to do this after the completion of any items that need to be completed and before free time—this can be.

Therapist should complete treatment integrity checklist for session 5—any items (e.g., items 1 and 2) not completed can be completed in a sixth session.

Appendix O: Programme Fidelity Checklist**Session 1**

_____A. Administer Pre-Treatment Assessment

- _____ 1. Therapist presented the rationale for the programme.
- _____ 2. Therapist encouraged questions.
- _____ 3. Therapist explained how anxiety shows itself. a. Body reactions b. Thinking c. Avoidance
- _____ 4. Therapist explained the concept of test anxiety.
- _____ 5. Therapist discussed the benefits of approach behavior.
- _____ 6. Therapist administered the test anxiety hierarchy to students.
- _____ 7. Therapist led discussion on hierarchy.
- _____ 8. Therapist noted how common and universal test/evaluation anxiety is.
- _____ 9. Students were taught how to check their pulses and the usefulness of this skill was discussed.
- _____ 10. Muscle relaxation exercises were demonstrated by therapist, and then students practiced on their own.
- _____ 11. Breathing exercises were demonstrated by therapist, and then students practiced on their own.
- _____ 12. Therapist discussed how these skills can be used to help students deal with test anxiety.

Session 2

- _____ 1. Therapist led a review discussion on how anxiety shows itself.
- _____ 2. Therapist noted how common and universal test/evaluation anxiety is.
- _____ 3. Basic study test preparation skills were discussed.
- _____ 4. Students discussed/participated.
- _____ 5. Therapist reviewed relaxation exercises and asked for feedback on their usefulness.
- _____ 6. Therapist taught “Stop, Drop, and Roll.”
- _____ 7. Students practiced relaxation techniques during imagined exposure to the lower level hierarchy items.
- _____ 8. Therapist discussed facing your fears (non avoidance of items on test anxiety hierarchy like avoiding studying).
- _____ 9. Therapist helped build student self-efficacy by reinforcing progress/learning.

Session 3

- ___ 1. Therapist led a review discussion on how anxiety shows itself.
- ___ 2. Therapist noted how common and universal test/evaluation anxiety is.
- ___ 3. Therapist reviewed relaxation exercises and asked for feedback on their usefulness.
- ___ 4. Students practiced relaxation techniques during imagined exposure to the remaining hierarchy items.
- ___ 5. Therapist discussed self-evaluation and self-reward/punishment.
- ___ 6. Students made a list of possible rewards for themselves.
- ___ 7. Students practiced creating hypothetical situations and demonstrated how they would evaluate and reward themselves.
- ___ 8. Students created personal success stories (e.g., drawing a self-portrait as a successful student; writing a short story about being a successful student).
- ___ 9. Therapist helped build student self-efficacy by reinforcing progress.

Session 4

- ___ 1. Students practiced relaxation techniques during imagined exposure to the remaining hierarchy items. Or Students practiced relaxation techniques during a mock test.
- ___ 2. Therapist reviewed previous information taught, including the “Stop, Drop, and Roll” technique and self-evaluations/self-rewards.
- ___ 3. Therapist led a review discussion on how anxiety shows itself.
- ___ 4. Therapist led a review discussion on progress.
- ___ 5. Therapist helped build student self-efficacy by reinforcing progress.

Session 5

- ___ 1. Therapist reviewed previous information taught, including the “Stop, Drop, and Roll” technique and self-evaluations/self-rewards.
- ___ 2. Students practiced relaxation techniques for remaining items on hierarchy and/or during a mock test.
- ___ 3. Therapist praised students’ progress and gave certificates of completion.
- ___ 4. Therapist administered post-intervention assessment (Unless there is another session planned if so administer at last session).

Session 6

As needed—what was completed?

Appendix P: Children's Test Anxiety Scale**TEST ATTITUDE SURVERY**

SAMPLE: Please read the following statement and decide if it best describes the way you are while you are taking tests. If the statement is **almost never** or **never** like you, you should circle **1**. If the statement describes how you are **some of the time**, circle **2**. If the statement describes the way you are **most of the time**, circle **3**. If the statement is **almost always** or **always** like you, circle **4**.

While I am doing tests....	ALMOST NEVER	SOME OF THE TIME	MOST OF THE TIME	ALMOST ALWAYS
I think about other things.	1	2	3	4

The rest of the items describe how some pupils may feel, think, or act while they are taking tests. Please read each statement carefully and decide if the statement describes how you think, feel or act during a test. Then circle the answer that best describes the way you are while taking a test. If you are not sure which answer to circle, read the statement again before circling your answer. Remember that there are no right or wrong answers on this survey. Please give truthful answers.

While I am doing tests....	ALMOST NEVER	SOME OF THE TIME	MOST OF THE TIME	ALMOST ALWAYS
1. I wonder if I will pass.	1	2	3	4
2. My heart beats fast.	1	2	3	4
3. I look around the room.	1	2	3	4
4. I feel nervous.	1	2	3	4
5. I think I am going to get a bad grade.	1	2	3	4

While I am doing tests....	ALMOST NEVER	SOME OF THE TIME	MOST OF THE TIME	ALMOST ALWAYS
6. It is hard for me to remember the answers.	1	2	3	4
7. I play with my pencil.	1	2	3	4
8. My face feels hot.	1	2	3	4
9. I worry about failing.	1	2	3	4
10. My belly feels funny.	1	2	3	4

While I am doing tests....	ALMOST NEVER	SOME OF THE TIME	MOST OF THE TIME	ALMOST ALWAYS
11. I worry about doing something wrong.	1	2	3	4
12. I check the time.	1	2	3	4
13. I think about what my grade will be.	1	2	3	4
14. I find it hard to sit still.	1	2	3	4
15. I wonder if my answers are right.	1	2	3	4

While I am doing tests....	ALMOST NEVER	SOME OF THE TIME	MOST OF THE TIME	ALMOST ALWAYS
16. I think I should have studied more.	1	2	3	4
17. My head hurts.	1	2	3	4
18. I look at other people.	1	2	3	4
19. I think most of my answers are wrong.	1	2	3	4
20. I feel warm.	1	2	3	4

While I am doing tests....	ALMOST NEVER	SOME OF THE TIME	MOST OF THE TIME	ALMOST ALWAYS
21. I worry about how hard the test is.	1	2	3	4
22. I try to finish up fast.	1	2	3	4
23. My hand shakes.	1	2	3	4
24. I think about what will happen if I fail.	1	2	3	4
25. I have to go to the bathroom.	1	2	3	4

While I am doing tests....	ALMOST NEVER	SOME OF THE TIME	MOST OF THE TIME	ALMOST ALWAYS
26. I tap my feet.	1	2	3	4
27. I think about how poorly I am doing.	1	2	3	4
28. I feel scared.	1	2	3	4
29. I worry about what my parents will say.	1	2	3	4
30. I stare.	1	2	3	4

Appendix R: Emotion Regulation Questionnaire for Children and Adolescents**ERQ-CA**

Below are a number of statements. Please read each statement, and then **circle the choice that seems most true for you**. Some of the statements may seem the same but they are different in important ways, so be sure to read carefully.

1. When I want to feel happier, I think about something different.	Strongly Disagree	Disagree	Half and half	Agree	Strongly Agree
2. I keep my feelings to myself	Strongly Disagree	Disagree	Half and half	Agree	Strongly Agree
3. When I want to feel less bad (e.g., sad, angry or worried), I think about something different.	Strongly Disagree	Disagree	Half and half	Agree	Strongly Agree
4. When I am feeling happy, I am careful not to show it.	Strongly Disagree	Disagree	Half and half	Agree	Strongly Agree
5. When I'm worried about something, I make myself think about it in a way that helps me feel better.	Strongly Disagree	Disagree	Half and half	Agree	Strongly Agree
6. I control my feelings by not showing them	Strongly Disagree	Disagree	Half and half	Agree	Strongly Agree
7. When I want to feel happier about something, I change the way I'm thinking about it.	Strongly Disagree	Disagree	Half and half	Agree	Strongly Agree
8. I control my feelings about things by changing the way I think about them.	Strongly Disagree	Disagree	Half and half	Agree	Strongly Agree
9. When I'm feeling bad (e.g., sad, angry, or worried), I'm careful not to show it.	Strongly Disagree	Disagree	Half and half	Agree	Strongly Agree
10. When I want to feel less bad (e.g., sad, angry, or worried) about something, I change the way I'm thinking about it.	Strongly Disagree	Disagree	Half and half	Agree	Strongly Agree

Appendix S: Study Limitations and Considerations for Future Research

Limitations	Future Recommendations
1. Non-active nature of control condition	Active control group to determine superiority of intervention
2. Limited follow-up time with high attrition rate in control group	Additional follow-ups (e.g. 6 months, 1 year, 2 years) Efforts to minimise attrition (e.g. phoning ahead of time to determine pupil presence)
3. Relatively small sample size	Increased sample size- to allow for increased chance of detecting significant effects, as well as examination of CTAS subscale scores (thoughts, behaviours, autonomic reactions), mediation analyses, and comparison of potential differences in outcomes based on demographic variables
4. Sample recruited from only one school	Recruit participants from several schools where possible to determine intervention effectiveness across school contexts and determine moderating effect of school cultures
5. Questionable reliability of two secondary outcome measures: cognitive reappraisal subscale on the ERQ-CA and Bandura's Self-Efficacy Scale for Academic Achievement	Further piloting of measures to determine reliability and adapting measures if necessary
6. Only one measure of primary outcome (TA)- due to parent attrition	Enhance the convenience of questionnaire completion (e.g. online) Information evening to enhance engagement, again supporting ease of attendance (e.g. online meeting forum or short video message) Highlight the importance of the research in this age group and the preventative benefits for those who it is not currently relevant to Alternatively, consider teacher reports for pupils high in TA or physiological measures of TA
7. Universal delivery contributing to a lack of relevance and potential for significant improvements –	Larger sample size to allow for categorisation of TA levels (i.e. low, medium, high) – to improve chance of detection of significance in those with higher levels of TA who the programme is more relevant to

- this impacted the overall mean change in TA scores
8. Implementation fidelity – self-report only
Include multi-informant measures of fidelity (i.e. self-report and direct observation) to increase confidence in conclusions regarding intervention outcomes
 9. Limited age range- fourth class only
Extend the participant sample to include older age groups with the purpose of improving generalisability of findings (e.g. fifth class and sixth class)
 10. Intervention delivery by researcher
Teacher-led delivery, supported by EPs, to determine effectiveness of this approach and support applicability in real-life setting
 11. Exclusively quantitative research methods- positivist paradigm
Employ a pragmatic paradigm- include both quantitative and qualitative data collection, (e.g. self-report measures and interviews to examine subjective experiences) to enhance the quality and depth of research findings
 12. Timing of intervention- contributed to reduced relevance
Conduct intervention in the lead up to standardised testing when it would be most relevant and beneficial to participants
 13. Hawthorne effect
Physiological measures of TA would provide an objective measure of TA

When using self-report measures, interpret with caution in light of potential Hawthorne effect
-