UNDERSTANDING THE PSYCHOLOGICAL MECHANISMS UNDERPINNING TEST ANXIETY

Thesis submitted in accordance with requirements of the University of Liverpool

for the degree of Doctor in Philosophy

by Christopher David Huntley

August 2021

Abstract

Test anxiety refers to concern about tests and the possible consequences of poor test performance, and is very common among university students. Worry, which refers to repetitive negative thinking, and emotionality, which refers to somatic anxiety experienced before and during tests, are the major dimensions of test anxiety. Worry is the cardinal feature of test anxiety as it much more strongly associated with reduced academic performance than emotionality.

This thesis begins with a meta-analysis of interventions for test-anxious university students to establish their relative efficacy compared to control conditions. Systematic searches located 44 eligible randomized controlled trials (RCTs). Interventions were moderately effective for reducing test anxiety and had weak effects for improving academic performance. There were no high quality RCTs. No intervention specifically targeted worry for therapeutic modification. Psychological models that can account for worry, and interventions based upon these models, may lead to better treatment outcomes.

Two psychological models of emotional disorder that can account for worry and emotionality are the Intolerance of Uncertainty (IU) and the Self-Regulatory Executive Function (S-REF) models. Each model proposes different mechanisms that lead to emotional disorder with negative beliefs about uncertainty emphasized in the IU model, and beliefs about thinking – metacognitive beliefs – emphasized in the S-REF model. Scant research has examined these models in test anxiety, and so a series linked empirical studies were conducted to examine their relative utility.

First, the psychometric properties of the principal measures of metacognitive beliefs and intolerance of uncertainty (IU), namely the

Metacognitions Questionnaire – 30 (MCQ-30) and Intolerance of Uncertainty Scale – 12 (IUS-12), were examined to assess their validity in UK university student samples. Factor analyses replicated the structures of the originally published MCQ-30 and IUS-12 providing support for their validity.

Next, two studies were conducted to test the utility of both models in test anxiety. First, the contributions of IU and metacognitive beliefs in explaining variance into test anxiety was examined in a cross-sectional study. Hierarchical linear regressions found metacognitive beliefs explained an additional 13% of variance in test anxiety, after controlling for IU. The final study, using a prospective study design, examined if: (1) IU and metacognitive beliefs predicted test anxiety, and (2) test anxiety mediated the relationship between test performance and IU/metacognitive beliefs. Only negative metacognitive beliefs about the uncontrollability and danger of worry predicted test anxiety, and the relationship between these beliefs and test performance was mediated by the worry, but not emotionality, dimension of test anxiety.

In summary, existing interventions are not particularly effective in reducing test anxiety and improving test performance. Examination of models found metacognitive beliefs are more important than IU in test anxiety, providing greater support for the S-REF model over the IU model. New interventions for test anxiety that can modify maladaptive metacognitive beliefs are needed. Reducing test anxiety will help students fulfill their academic potential.

Table of Contents

Chapter One: Introduction to test anxiety

Chapter Two (Study 1): Efficacy of interventions for test anxious university students: A meta-analysis of randomized controlled trials

Chapter Four (Study 2): Assessing metacognitive beliefs and intolerance of uncertainty among university students: Psychometric properties of the Metacognitions Questionnaire – 30 (MCQ-30) and Intolerance of Uncertainty Scale – 12 (IUS-12)

Chapter Five (Study 3): The relative contributions of intolerance of uncertainty and metacognitive beliefs to test anxiety

Chapter Six (Study 4): Do metacognitive beliefs and intolerance of uncertainty prospectively predict test anxiety, and does test anxiety mediate their relationship to test performance?

Chapter Seven: General discussion and conclusions

List of figures

List of Tables

List of Appendices

Glossary

Abbreviation	Definition
AP	Analytical Psychotherapy
BCa	Bias corrected and accelerated
BT	Behaviour Therapy
CAS	Cognitive Attentional Syndrome
CFI	Comparative Fit Index
CBT	Cognitive-Behavioural Therapy
CFA	Confirmatory Factor Analysis
CI	Confidence Interval
CONSORT	Consolidated Standards of Reporting Trials
EFA	Exploratory Factor Analysis
GAD	Generalized Anxiety Disorder
GDT	Guideline Development Tool
GRADE	Grading of Recommendations Assessment, Development and Evaluation
HTA	High-test-anxious
GPA	Grade Point Average
GRADE	Grading of Recommendations, Assessment, Development, and Evaluations
IU	Intolerance of Uncertainty
IUS-12	Intolerance of Uncertainty Scale – 12
IU-I	Inhibitory IU
IU-P	Prospective IU

LTA	Low-test-anxious
M	Mean
MCQ-30	Metacognitions Questionnaire – 30
MCQ-30-CC	MCQ-30 Cognitive Confidence
MCQ-30-CSC	Cognitive Self-Consciousness
MCQ-30-NC	Need to Control Thoughts
MCQ-30-NEG	Negative beliefs about uncontrollability and danger of worry
MCQ-30-POS	Positive beliefs about worry
MCT	Metacognitive Therapy
MDD	Major Depressive Disorder
MRC	Medical Research Council
MTA	Moderately-test-anxious
NPT	Normalisation process theory
OCD	Obsessive-Compulsive Disorder
OSCE	Objective Structured Clinical Examination
PCC	Person-Centred Counselling
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PTSD	Post-Traumatic Stress Disorder
RCT	Randomized Controlled Trial
REBT	Rational Emotive Behavioural Therapy
RMSEA	Root Mean Square Error of Approximation
SAD	Social Anxiety Disorder
SE	Standard Error

SD	Standard deviation
SEM	Structural Equation Modeling
SMD	Standardized Mean Difference
S-REF	Self-Regulatory Executive Function (model)
SPIRIT	Standard Protocol Items: Recommendations for Interventional Trials
SPSS	Statistical Package for the Social Sciences
SRMR	Standardized Root Mean Square Residual
SST	Study Skills Training
STICSA-S	State-Trait Inventory of Cognitive and Somatic Anxiety – State Scale
TAI	Test Anxiety Inventory
WLSMV	Weighted Least Squares Mean and Variance adjusted estimator
WRMR	Weighted Root Mean square Residual

Acknowledgements

I would like to thank the following people for helping me to complete my thesis. First, thank you to my supervisors, Dr Peter Fisher, Professor Bridget Young, and Professor Catrin Tudur Smith who have supported me, challenged me, and provided constructive criticism throughout my PhD. I am particularly grateful to my primary supervisor, Dr Fisher, for his thought provoking and often provocative questions and words of wisdom ("consistency of language!"). I would also like to thank Professor Vikram Jha, who initially supervised me before leaving academia.

Additionally, I would like to thank Dr James Temple and Dr Melissa Longworth who assisted in screening and data extraction in Study 1.

Lastly, I would like to thank my wife, Claire Huntley for being my rock throughout my PhD studies. We have been through a lot since I started this PhD, including having another child, who was subsequently very ill but has thankfully pulled through and is now flourishing, and two house moves (not to mention COVID-19). I am looking forward to a less frantic future!

Thank you all for your support and wisdom.

Published and Submitted Publications from Thesis

Huntley, C. D., Young, B., Jha, V., & Fisher, P. L. (2016). The efficacy of interventions for test anxiety in university students: A protocol for a systematic review and meta-analysis. *International Journal of Educational Research, 77*, 92-98. doi:10.1016/j.ijer.2016.03.001

Huntley, C. D., Young, B., Temple, J., Longworth, M., Smith, C. T., Jha, V., & Fisher, P. L. (2019). The efficacy of interventions for test-anxious university students: A meta-analysis of randomized controlled trials. *Journal of Anxiety Disorders, 63*, 36-50. doi:10.1016/j.janxdis.2019.01.007

Huntley, C. D., Young, B., Tudur Smith, C., Jha, V., & Fisher, P. L. (2020). Assessing metacognitive beliefs in test anxiety: Psychometric properties of the metacognitions questionnaire – 30 (MCQ-30) among university students. *Current Psychology*. doi:10.1007/s12144-020-00662-y

Huntley, C. D., Young, B., Tudur Smith, C., & Fisher, P. L. (2020). Uncertainty and test anxiety: Psychometric properties of the Intolerance of Uncertainty Scale – 12 (IUS-12) among university students. *International Journal of Educational Research*. doi: 10.1016/j.ijer.2020.101672

Huntley, C. D., Young, B., Tudur Smith, C., Jha, V., & Fisher, P. L. (submitted). Testing times: The contribution of intolerance of uncertainty and metacognitive beliefs to test anxiety in college students. *BMC Psychology*.

Huntley, C. D., Young, B., Tudur Smith, C., Jha, V., & Fisher, P. L. (in prep). Test anxious medical students need to change how they think rather than what they think.

Declaration

I, Christopher Huntley, declare that I am the author of this thesis, that, unless otherwise stated, all references cited have been consulted by me, that unless otherwise stated, the work of which this thesis is a record has been done by myself and has not been previously accepted for a higher degree.

Christopher Huntley

August 2021

Thesis Overview

Test anxiety concerns the fear of tests and performance-evaluative situations. Test anxiety undermines test and academic performance, meaning students' potential can go unfulfilled. It is therefore critical to identify and treat test anxiety.

This thesis has two broad aims: (i) to establish the efficacy of existing interventions for test anxious university undergraduates, and (ii) to understand the psychological beliefs and processes that underlie test anxiety. Four linked studies were conducted to address these aims.

Chapter 1 provides an overview of test anxiety, its impact on test and academic performance, and a brief description of existing psychological theories of test anxiety.

Chapter 2 first presents an overview of interventions for test anxiety and an evaluation of past meta-analytic reviews of these interventions, concluding that the efficacy of interventions for test anxious university students is unknown. Subsequently, a meta-analysis of randomized controlled trials examining interventions for test anxious university students (Study 1) is presented, with the aim of establishing the efficacy of interventions for (i) reducing test anxiety, and (ii) improving test and academic performance. The review protocol and the final meta-analysis were both published (Huntley et al., 2016; Huntley et al., 2019).

Chapter 3 provides an overview of two contemporary theoretical models of anxiety, the Intolerance of Uncertainty (IU) model (Dugas, Gagnon, Ladouceur, & Freeston, 1998) and Self-Regulatory Executive Function (S-REF)

model (Wells & Matthews, 1994, 1996), which can account for the central feature of test anxiety, namely worry.

Prior to evaluating the utility of the IU and S-REF models in test anxiety, it was first necessary to establish the validity of the most commonly used measurement tools for assessing beliefs considered central to each model. Therefore, Chapter 4 examines the factorial validity of the Intolerance of Uncertainty Scale – 12 (IUS-12) and the Metacognitions Questionnaire – 30 (MCQ-30) among UK university students, using factor analysis, and examines relationships between these constructs and test anxiety using structural equation modeling (Study 2). Papers reporting the psychometric properties of the MCQ-30 and IUS-12 were published (Huntley, Young, Tudur Smith, Jha, & Fisher, 2020; Huntley, Young, Tudur Smith, & Fisher, 2020).

Next, the utility of the IU and S-REF models in test anxiety are examined in cross-sectional (Chapter 5: Study 3) and prospective (Chapter 6: Study 4) studies. Both IU and metacognitive beliefs are expected to be associated with test anxiety. However, the S-REF model proposes that beliefs in the metacognitive domain, which are involved in the monitoring, appraisal, and regulation of cognitions, are more important in the etiology of emotional disorder than beliefs in the cognitive domain, such as IU. This informed the main prediction; metacognitive beliefs would explain additional variance in test anxiety after controlling for IU. Hierarchical multiple linear regressions were used to test this main prediction. Finally, the relationships between metacognitive beliefs, test anxiety, and examination were explored. Papers reporting these studies have been submitted to peer-reviewed journals.

Finally, Chapter 7 summarizes the main findings of the thesis, study limitations, practical implications of this research, and suggestions for future research.

Chapter One

Introduction

"*Test anxiety is the interest paid on academic troubles before it is due.*"

Covington & Omelich (1988)

1.1 Introduction

Anxiety about examinations, or *test anxiety*, has been a major educational concern for over a century. Test scores are important in determining the progression, prospects, and mental wellbeing of students. Test anxiety is a pervasive problem in higher education settings as it undermines performance and academic achievement (Howard, 2020; Spielberger & Vagg, 1995; Zeidner, 1998). It is therefore important to identify, understand, and alleviate test anxiety to help students fulfil their academic potential.

1.2 What is test anxiety?

Test anxiety is a situation-specific form of anxiety consisting of the cognitive, affective, and behavioural responses to fear of tests or performance-evaluative situations (Speilberger & Vagg, 1995; Zeidner, 1998). Performance-evaluative situations are ones in which an individual is assessed, implicitly or explicitly, against a standard of performance. An example of a performance-evaluative situation is when a student is asked to demonstrate a technique, such as solving a mathematics equation, in front of the class. Tests or examinations are a special form of performance-evaluative situation. Here, individuals are required to complete set tasks under circumscribed conditions and are evaluated against external criteria. Tests have important academic consequences such as determining whether a student can progress into the next academic year.

Highly test-anxious students have a low threshold for experiencing anxiety in tests or performance-evaluative situations. Such situations are viewed as personally threatening (Zeidner, 1998), coupled with concern about embarrassment of failing (Sarason & Sarason, 1990). The degree of perceived

threat is determined by: (i) salience of the test to the individual (e.g., “I need to pass this test to progress”), (ii) the perceived negative outcomes of the test (e.g., “I will not to be able to apply for postgraduate programmes if I do not pass”), (iii) imminence of the test (e.g., “The test is next week I have not started studying”), (iv) the perceived aversiveness of the test (e.g., “I hate doing statistics tests!”), and (v) perceived coping capabilities (e.g., “I don’t even know where to start studying for this test”) (Zeidner, 1998).

Test anxiety does not constitute a separate diagnostic category in the fifth edition American Psychiatric Association’s (APA) Diagnostic and Statistical Manual of Mental Disorders (DSM-5; APA, 2013). However, prior to the publication of DSM-5, test anxiety was considered for inclusion as a specific phobia subtype (LeBeau et al., 2010). However, its inclusion was rejected owing to the lack of epidemiological data, and because test anxiety symptomology extends beyond an irrational fear of a specific situation to also encompass fears about negative evaluation (LeBeau et al., 2010). The lack of diagnostic status has meant there is no universally agreed way of identifying clinically or highly test anxious individuals. Rather than meeting diagnostic criteria, students are typically classified as low-test-anxious, moderate-test-anxious, and high-test-anxious (e.g., Cassady & Finch, 2015; von der Embse et al., 2014).

1.2.1 The major dimensions of test anxiety

Worry and emotionality are widely accepted as the major dimensions of test anxiety. Worry refers to repetitive negative thinking that is primarily verbal-conceptual and future-oriented in nature (Borkovec et al., 1983). In test anxiety, worry primarily concerns the potential consequences of failing a test (Liebert &

Morris, 1967; Tempel & Neumann, 2016). Emotionality refers to the physiological arousal and somatic symptoms experienced in test contexts, such as heart-rate accelerations, muscle tension, and excessive sweating.

Worry is considered the central feature of test anxiety as it is more strongly linked to poorer test and academic performance than emotionality (Cassady, 2002; Sarason, 1988, Wine, 1971). Highly test anxious students become self-occupied with the potential negative consequences of failing, spending much of their time worrying during tests (Deffenbacher, 1978). Persistent worrying is associated with negative affect (Borkovec et al., 1983; Newman et al., 2019), and high-test-anxious individuals report more negative thoughts than their peers (Galassi, Fierson, & Sharer, 1981).

Worry can used as a form of problem solving and an attempt to take control of emotions and arousal (Borkovec et al., 1983). However, pathological or persistent worry can become a cognitive avoidance strategy aimed at reducing physiological arousal experienced when feeling emotional (Borkovec, et al., 1983; Eysenck, 1992). Worry therefore persists as it is negatively reinforced due to the dampening effect it has on the anxiety-related physiological arousal (Borkovec et al., 1983). High-test-anxious individuals report persistence of worry before, during, and after tests (Kim & Rocklin, 1994). Ultimately, worry has deleterious effects on self-regulation because it results in the inhibition or avoidance of emotional processing (Borkovec & Hu, 1990; Borkovec et al., 1993; Newman et al., 2004). Periods of worry lead to greater anxiety, less task-focused attention, and more negative thoughts (Borkovec et al., 1983).

Test anxiety is associated with actual elevated physiological arousal (Roos, et al., 2021). However, high- and low-test-anxious individuals cannot be

reliably differentiated based on indices of physiological arousal, such as heart rate and skin conductance (Hollandsworth et al., 1979; Holroyd et al., 1978). Therefore, emotionality refers to an individual's perceived physiological arousal rather than the actual arousal (Schwarzer, 1984). High-test-anxious individuals perceive their arousal as being debilitating, whereas low-test-anxious individuals take their arousal as a call to action (Hollandsworth et al., 1979). Emotionality typically peaks at the beginning of the test before steadily decreasing (Lieberts & Moms, 1967; Kim & Rocklin, 1994).

High-test-anxious students are also characterized by several behavioural deficits, particularly around study skills, such as taking and organizing notes, avoidance of studying, and usage of time in tests (Culler & Holahan, 1980; Kirkland & Hollandsworth, 1979, 1980). For example, Culler and Holahan (1980) found high-test-anxious students scored a standard deviation below low-test-anxious students on their study skill habits despite reporting spending more time studying. Avoidance is common in anxiety problems (Hofmann & Hay, 2018) and test anxious individuals report wishing to avoid or escape test situations (Galessi, Frierson, & Sharer, 1981a,b). However, individuals are acutely aware that avoiding tests is not a realistic option, and therefore use worry as a cognitive avoidance strategy instead (Carver, 1996). In experimental studies where participants can voluntarily disengage from a difficult task, high-test-anxious students disengaged significantly earlier than their less test anxious peers (e.g., Geen, 1987). Test anxiety is associated with avoidant-based coping strategies (Putwain et al., 2016; Stober, 2014).

1.2.2 State and trait test anxiety

Trait test anxiety refers to relatively unchanging and enduring negative reactions to tests. Trait test anxiety scores have been shown to be stable over a three-month period (Hong, 1998).

State test anxiety refers to the momentary experience of anxiety in test contexts, i.e., individuals are typically asked how they feel at the moment (or last few minutes) rather than how they typically feel. The degree of state test anxiety experienced is the product of individual personality differences (e.g., trait test anxiety) and the specific test context. State test anxiety increases as test approaches, reaching its peak at the beginning of a test (Doctor & Altman, 1969; Lotz & Sparfeldt, 2017).

1.3 Origins of test anxiety

Test anxiety starts to manifest itself from primary or elementary school onwards (Hembree, 1988; von der Embse et al., 2018). Aside from the general propensity to experience anxiety that has biological or genetic origins, two primary antecedents specific to test anxiety have been proposed. Firstly, it is thought that interpretation of tests as potentially threatening is based upon an individual's negative historical experiences with tests (Sieber, 1980 cf. Ziedner, 1998). Behavioural conceptualizations emphasize that the test becomes a conditioned or aversive stimulus if paired with an aversive event, such as negative feedback from a teacher or social shaming by peers (Garlington & Carter, 1968). Secondly, the child-parent/carer relationship is also proposed to be an important factor in the development of test anxiety. Excessively high expected

standards of parents may induce anxiety in their child who struggles to meet this high bar (Teichman & Ziv, 1992; Krohne, 1992).

1.4 How is test anxiety assessed?

A great many self-report questionnaires assessing test anxiety have been developed. Most questionnaires developed assess trait test anxiety. Consistent with initial theoretical thinking in the field, early questionnaires such as the Test Anxiety Questionnaire (TAQ; Mandler & Sarason, 1952) assessed test anxiety purely as a unidimensional construct (Sarason & Mandler, 1952). Following Liebert and Morris' (1967) distinction of worry and emotionality as the major dimensions of test anxiety, many questionnaires have been developed assessing these dimensions, such as the Worry/Emotionality Questionnaire (WEQ; Liebert & Morris, 1967) and Test Anxiety Inventory (TAI; Spielberg, 1980).

Several questionnaires have assessed dimensions of test anxiety beyond worry and emotionality. For example, the Reaction-to-Tests (RTT; Sarason, 1988) assesses two purported extra dimensions of test anxiety, labelled 'test-irrelevant thinking' and 'tension', whilst others have reported the additional components of 'cognitive interference' (Friedman & Bendas-Jacob, 1997) and 'distraction' (Hodapp & Benson, 1997). However, there is weaker support for the reliability and validity of these and other more differentiated dimensions, a lack of consistency in the constructs they are purported to tap into which limits the interpretability of the subscale scores, and researchers have not universally accepted or used these other dimensions (Anderson & Sauser, 1995; Spielberger & Vagg, 1995; von der Embse et al., 2018). Most recently, a new questionnaire assessing trait test anxiety, the Multidimensional Test Anxiety Scale (MTAS;

Putwain et al., 2021; von der Embse, Putwain, & Francis, 2021) has been developed, with item content based upon past test anxiety scales, psychological theories of test anxiety, and a panel of test anxiety experts. The MTAS consists of four correlated factors, assessing two cognitive dimensions ('worry', 'cognitive interference') and two affective-physiological arousal dimensions ('tension', 'psychological indicators') of trait test anxiety. Initial examination of the MTAS' psychometric properties found support for both correlated factors (i.e., four correlated factors) and second order (i.e., a trait test anxiety factor that accounts for the four sub-factors) measurement models. The MTAS also showed convergent validity with TAI, good test-retest reliability over a four-month period (r = .80), and good internal consistency (Putwain et al., 2021).

Of the trait test anxiety questionnaires, the TAI has the most significant support for its reliability and validity (Chapell et al., 2005; Everson, Millsap, & Rodriguez, 1991; Spielberger, 1980; Szafrankski, Barrera, & Norton, 2012; Xu, Cai, & Tu, 2020). The TAI consists of a strong general factor and reliable and interpretable correlated worry and emotionality dimensions (Xu, Cai, & Tu, 2020). In practice, this means either TAI total or subscale scores can be used in research and educational settings (Xu, Cai, & Tu, 2020).

There are fewer questionnaires assessing state test anxiety, with only the Pre-Examination Questionnaire (PEQ; Liebert & Morris, 1967) and its revision, the Worry-Emotionality Questionnaire (WEQ; Morris et al., 1981) available. Both the PEQ and WEQ lack empirical support for their validity, with only evidence for their internal reliability available (Liebert & Morris, 1967; Morris et al., 1981). Given the aim is to measure anxiety in the moment, or at least a very prescribed period (e.g., past 10 minutes), researchers can use more general

questionnaires of state anxiety in test contexts. The State-Trait Anxiety Inventory – State (STAI-S; Spielberger et al., 1983) has been most frequently used in test anxiety (Hembree, 1988; Gros et al., 2007). However, the STAI-S also assesses mood and excitability in addition to anxiety (Bieling et al., 1998). This has led to newer and purer measures of state anxiety being developed including the STICSA-S (Gros et al., 2007; Ree et al., 2008) and SSSQ (Helton 2004; Matthews et al., 2002). Both the STICSA-S and SSSQ assess state cognitive/worry and somatic dimensions of anxiety (i.e., emotionality). At present, only the STICSA-S has been validated in test contexts (Ree et al., 2008).

1.5 Prevalence of test anxiety

University students frequently report test anxiety as a primary concern in academic surveys (Knappe et al., 2011; Spielberger, Anton & Bedell, 1976). Precise estimates of test anxiety prevalence have been difficult to establish due to definitional (i.e., lacking a consistent operational definition of test anxiety) and methodological reasons (i.e., no gold standard measure) (King & Ollendick, 1989). However, based on observational data approximately 20-25% of university students are highly test anxious (Hill & Wigfield, 1984; Naveh-Benjamin et al., 1997; Neurderth, Jabs, & Schmidtke, 2009; Thomas, Cassady, & Finch, 2018). Supporting biochemical evidence for test anxiety prevalence came from the first known study investigating test anxiety (Folin, Demis, & Smillie, 1914; cf. Spielberger & Vagg, 1995). This study found around 20% of students had glycosuria (i.e., excess sugar in the urine, and evidence of anxiety), after taking a test, when no student showed any evidence of glycosuria before the test.

1.6 Demographics differences in test anxiety

More females report experiencing test anxiety than males, and females also report greater test anxiety severity than males (Hembre, 1988; Putwain & Daly, 2014; Seipp & Schwarzer, 1996; von der Embse et al., 2018). Gender group differences in severity is a result of differences in scores for emotionality but not worry, with females reporting greater emotionality than males (Deffenbacher, 1980; Zeidner & Nevo, 1993). However, measurement of actual physiological indices (e.g., pulse-rate) reveal no gender differences (Deffenbacher, 1986).

Although it is speculated that older adults may be more anxious in test situations than younger adult (Ross, 1968), there is no empirical data that supports this contention. Test anxiety prevalence remains relatively constant across age ranges from 9-10 years old and onwards (Hembree, 1988).

Test anxiety is reported by all ethnic groups (Hembree, 1988). People of non-European/white heritage report greater test anxiety than their European/white heritage peers in both community samples (Thames et al., 2015) and pre-secondary education (Clawson et al., 1989; Payne, 1984; Putwain, 2007; Wren & Benson, 2004). Test anxiety is experienced at similar severity levels across countries with different cultural backgrounds (Seipp & Schwarzer, 1996).

Socio-economic status (SES), which refers to family economic and social status (e.g., parental income and professional status), is generally significantly negatively associated with test anxiety among high school students, i.e., higher SES is associated with lower test anxiety (Guida & Ludlow, 1989; Putwain, 2007, 2008; Xu et al., 2021). However, some studies have found no significant relationships or even found higher SES is associated with greater test anxiety

(Chen, 2012; Chukwuorji & Nwonyi, 2015). No study has investigated the relationships between SES and test anxiety among university students. However, first-generation university students (i.e., those who parents and family have not been to university), who are more likely to come from lower SES backgrounds (Wilbur & Roscigno, 2016), experience greater test anxiety than continuing-generation students (i.e., those with at least one parent who went to university) (Janke et al., 2017).

1.7 Individual differences in test anxiety

A growing range of individual differences have been linked to test anxiety. Low self-efficacy or perceived low academic confidence is associated with test anxiety (Pekrun, 2006; Pekrun et al., 2004), whilst high confidence in one's ability to perform well in a test is negatively associated with test anxiety (Reeve et al., 2008; Zohar et al., 1998).

Academic self-concept, defined as an internal locus of control over academic tasks, is negatively associated with test anxiety (von der Embse et al., 2018). High academic self-esteem, defined as beliefs in one's ability to perform academic tasks competently, is also negatively associated with test anxiety (von der Embse et al., 2018).

Motivation is consistently associated with academic achievement (Cerasoli, Nicklin & Ford, 2014). Although there are several theories of motivation (Olsson, 2008), a common means of investigating motivation has divided it into two categories (Ryan & Deci, 2000): extrinsic motivation, where external factors (e.g., praise, money) drive reward behaviour and, intrinsic motivation, where internal factors (e.g., the individual finds the task satisfying)

drive reward behaviour. Test anxiety is positively associated with extrinsically motivated behaviour and negatively associated with intrinsically motivated behaviour (von der Embse et al., 2018).

Test anxious students engage in different ways of coping. The two most examined ways of coping are problem-focused and avoidance-based coping. Problem-focused coping refers to attempts to manage or solve the problem by removing, circumventing, or ameliorating the stressor, while avoidance coping refers to engaging in non-task related thinking or behaviours to avoid the stressor (Endler & Parker, 1990). Problem-based coping is weakly negatively associated with test anxiety, but avoidance-based coping is strongly positively associated with test anxiety (Matthews, Hillyard, & Campbell, 1999; von der Embse et al., 2018).

The relationships between the "big five" personality factors and test anxiety have been explored in many studies. The big five personality factors consist of: (i) neuroticism, which refers to the tendency to experience high levels of nervousness and apprehension, (ii) extraversion, that concerns to how outgoing an individual is, (iii) openness which refers to curiosity, imagination, and openness to new experiences, (iv) conscientiousness that concerns thoughtful and organized goal-oriented behaviour, and (v) agreeableness which refers to how cooperative and prosocial an individual is. Meta-analytic findings show neuroticism is strongly positively associated with test anxiety (r = .46), conscientiousness is negatively associated (r = -.18), while there are weak or non-significant relationships openness, extraversion, and agreeableness (von der Embse et al., 2018).

Student beliefs and reasons for task engagement can impact upon academic performance (Pintrich, 2007). Elliot and McGregor (1999) proposed a 2 x 2 model with mastery (i.e., desire to master skills for personal development) and performance (i.e., demonstrate competence relative to others) goals interacting with approach (i.e., wish to do well) and avoidance (i.e., wish to avoid failure) dimensions. Four achievement goals are articulated: (i) mastery-approach represents the goal of obtaining task competence, (ii) mastery-avoidance represents the goal of avoiding incompetent task performance, (iii) performance-approach represents the goal of performing better than some peer-referenced norm, and (iv) performance-avoidance represents the goal of avoiding worse than some peer-referenced norm. There are medium associations between test anxiety and mastery and performance avoidance goals (*r*s of .30 and .37 respectively). There is a positive, albeit small, relationship between performance-approach and test anxiety (r = .09), but no significant relationship between mastery-approach and test anxiety (von der Embse et al., 2018).

Procrastination and perfection have been linked to test anxiety. Academic procrastination is defined as the purposeful delaying of starting or completing tasks for irrational reasons (Tibbett & Ferrari, 2019). Perfectionism refers to the combination of extremely high personal standards and excessively critical self-evaluations (Burgess, Frost, & DiBartolo, 2016). Meta-analyses have found procrastination and perfectionism are significantly and positively associated with test anxiety (Burcas & Cretu, 2021; Steel, 2007; Van Eerde, 2003).

1.8 Situational determinants of test anxiety

Test parameters can moderate the degree of test anxiety experienced by students. Examples of test parameters include test difficulty (or at least perceived difficulty), stakes, novelty, time pressures, and environmental cues. Perceptions of a test as difficult heightens test anxiety (Hembree, 1988; Pekrun et al., 2004; von der Embse et al., 2018). High stakes tests are associated with greater test anxiety than low stakes tests (Segool et al., 2013; Silaj et al., 2021; Reeve, Bonaccio, & Charles, 2008; von der Embse et al., 2018). Familiarity with test and context can reduce test anxiety (Anstasi, 1981; von der Embse et al., 2018). In evaluative, but not test conditions, high-test-anxious students reported greater worry than their peers (Deffenebacher, 1978; Zatz & Chassin, 1985). The order of examination questions can affect the degree of test anxiety experienced by an individual, with high-test-anxious individuals reporting greater anxiety if they are presented with difficult items at the start of an examination (Zeidner, 1998).

Students find problem solving and essay-style assessment more anxiety provoking than multiple choice questions (Green, 1981). Students in clinical undergraduate degree programmes (e.g., medicine) report Objective Structured Clinical Examinations (OSCE), an examination format in which students perform clinical tasks in front of examiners, the most anxiety provoking of all examination formats (Guraya et al., 2018). In the general test anxiety literature, the presence of an external observer, is more anxiety provoking (Geen, 1976, 1985), and this may explain the increased anxiety felt by students in OSCEs.

1.9 Test anxiety and academic performance

Many studies have found test anxiety is associated with poorer examination performance. Meta-analyses of these studies consistently conclude that test anxiety is associated with poorer examination and academic performance (Hembree, 1988; Seipp, 1991; von der Embse et al., 2018).

The first two meta-analyses, synthesized data across all ages and educational levels (i.e., from primary/elementary school level to higher education) (Hembre, 1988; Seipp, 1991). Test anxiety and academic performance were significantly negatively correlated (*r*s from -0.23 to -0.29; Hembree, 1988; Seipp, 1991), with worry more strongly correlated with poor performance (*rs* around –0.31) than Emotionality (*rs* around –0.15). Significant heterogeneity within worry but not emotionality suggests other person-situation variables influence the relationship between worry and academic performance (Seipp, 1991).

The most recent meta-analysis (von der Embse et al., 2018) analysed data from 238 studies published between 1988 and 2018. Rather than aggregating across all age groups, data were analysed for different age and educational groups. Test anxiety in university students (53 studies; n = 20,849) was associated with poorer academic performance (r = –0.27). Sub-analyses examining worry and emotionality dimensions synthesized data from across all age and educational levels, with worry (95% CIs -.31 to -.28) again more strongly associated with test anxiety than emotionality (95% CIs -.20 to -.16).

Females report greater test anxiety than males but they do not perform worse in tests (Hembree, 1988). This is likely because females report greater

emotionality but not worry, and worry is the primary factor in reduced examination performance.

1.9.1 How does test anxiety affect test performance?

There are two accounts for the deleterious effects of test anxiety on test performance: interference and deficit accounts.

Interference accounts are situated within the broader information processing models dominant in cognitive psychology (Baddeley, 2012; Moran, 2016). These accounts propose that test anxiety has a casual effect on test performance, primarily due to the interference of worry and attention on cognitive processes (Derakshan & Eysenck, 2009; Eysenck & Calvo, 1992; Eysenck et al., 2007; Sarason, 1984; Wine 1971). Test-anxious individuals divide their attention and mental resources between self-occupied and task-based thinking (Eysenck et al., 2007; Sarason, 1980a; Wine 1971). Anxiety reduces performance efficiency by utilizing the same central executive and memory resources that would be better deployed to the task-at-hand (Derakshan & Eyseck, 2009, Derakshan, Smyth & Eysenck, 2009; Eysenck & Byrne, 1994; Hopko, Hunt, & Armento, 2005).

Anxiety may not always negatively impact performance. Current theorizing has highlighted the distinction between processing efficiency and processing effectiveness. Processing efficiency refers to the degree of mental resources used to complete a task, while processing effectiveness refers to performance outcomes (Eysenck et al., 2007). Anxious individuals employ strategies to try and mitigate the adverse effects of anxiety, such as increased effort (Ikeda, Iwanage, & Seiwa, 1996). Test anxiety therefore impacts

performance efficiency greater than performance effectiveness and will have the most debilitating effects when anxiety is high (e.g., because a test perceived as very important) and the task is complex and difficult (e.g., involves a high cognitive 'load').

The impact of anxiety, and worry, has been confirmed in experimental studies (Angledis et al., 2019), with mental processing particularly impacted by disruptions in inhibition (i.e., ability to keep focus on the task) and switching (i.e., flexible and managed deployment of attention) functions in highly anxious individuals (Darvishzadeh, Aguilar, & Moradi, 2012; Derakshan & Eysenck, 2009; Eysenck & Derakshan, 2011; Eysenck et al., 2007).

Contrary to interference account, deficit accounts propose that test anxiety is correlated but not casually linked to test performance (Culler and Holahan, 1980; Tobias, 1985; Reeve & Bonaccio, & Charles 2008; Sommer & Arendasy, 2014). Here, test anxiety is merely a concomitant of a test-taker's awareness of their lack of preparedness for the test (Paulman & Kennelly, 1984; Smith, Snyder, & Handelsman, 1982). Deficits can manifest in specific domains, such as study skills (e.g., Klinger, 1984) or in more general cognitive abilities such as working memory (e.g., Bishop, 2009). There is contradictory evidence that high-test anxious students have poorer study and test-taking habits than their low-test-anxious peers, with some studies finding significant differences (Bruch, 1981; Kirkland & Hollandsworth, 1980; Paulman & Kennelly, 1984) and others finding no differences (Bruch, Juster, Kaflowitz, 1983; Culler & Holohan, 1980). Researchers have also noted that anxiety can develop independently of study skills, such that an individual may have good study skills but still feel anxious, and vice versa (Zeidner, 1998). Moreover, though deficit models propose

inadequate study skills result in test anxiety, structural equation modelling has shown that test anxiety influences study style, with high-test-anxious individuals more likely to adopt surface (e.g., re-reading), as opposed to deep (e.g., interrogation and integration of material), learning approaches (Spada & Moneta, 2012, 2014).

Overall, there is much more empirical support for interference accounts of test anxiety on test performance over the deficit accounts. Interference accounts, which have strong cognitive psychology underpinnings, can explain how anxiety and worry interfere with central executive functions (e.g., attentional control) and working memory to reduce test performance.

1.10 Test anxiety and mental wellbeing

Test anxiety is associated with poorer student mental health. Highly test anxious students report poorer mental health (Depreeuw & DeNeve, 1992) and are more likely to dropout or repeat a year of study (Schaefer et al., 2007 cf. Neuderth, Jabs, & Schmidtke, 2009) than their low-test-anxious peers.

Students with mental health issues are more likely to have alcohol related issues than their peers (Markman et al., 2004). Test anxiety moderates the relationship between protective behavioural strategies regarding alcohol consumption (e.g., avoiding attempts to keep up or out drink others) and alcohol-related problems (Linden, Lau-Barraco, & Milletich, 2012; Tektas, Paulsen, & Sel, 2013). High-test-anxious students report greater use of non-prescription stimulants for non-medical reasons (Ne'Eman-Haviv & Bonny-Noach, 2019; Sattler & Wiegel, 2013). A US survey found 8.3% of 4,580 of college undergraduates engaged in illicit prescription drug use for the purposes of

‘cognitive enhancement’ in their lifetime, and 5.9% within the last year (McCabe et al., 2006, 2008). Increased usage of illicit prescription drug use for cognitive enhancement is associated with test anxiety (Sattler & Wiegel, 2019).

Test anxiety is a potential risk factor for psychopathology in university students (Tosevski, Milovancevic, & Gajic, 2010). Test anxiety is particularly linked to social anxiety (Herzer, Wendt, Hamn, 2014; Kavakci, Guler, & Cetinkaya, 2014). Test anxious individuals report high fears of negative social evaluation (Lowe & Lee, 2008), and 54% of test-anxious children also met criteria for an anxiety disorder, with social anxiety disorder most frequently comorbid (Biedel et al., 1994; Biedel & Turner, 1988). However, no differences have been found in social anxiety scores between high- and low-test anxious students (Hall, 2005).

Depressive symptomology, low negative affect and avoidant behaviour are associated with test anxiety (Akinsola & Nwajei, 2013; Huberty & Dick, 2006), as is suicidal ideation (Lee et al., 2006). Repeat test-takers report significantly higher suicidal ideation than first-time test-takers (Mamun et al., 2021). Academic pressures are an important factor in suicides amongst those in education, with 29% of suicide cases involving students facing test or test results at the time of death (Rodway et al., 2016).

Test anxiety is one of the principal reasons for students seeking access to mental health and university student support services (Rückert, 2015). In a national survey, 13% of German university students sought counselling for test anxiety (Middendorf et al., 2016 cf. Schillinger et al., 2021). The number of support-seeking students for test anxiety is likely underestimated, as 11.3% of 1,099 students surveyed at a Canadian university stated that they would not seek

help for their test anxiety (Gerwing et al., 2015). Not seeking help was due to not wishing to be seen as 'weak' and believing that academic staff would be unable or unwilling to help (Gerwing et al., 2015).

1.11 Psychological theories of test anxiety

Systematic investigation of test anxiety began in the 1950s (Mandler & Sarason, 1952) and theoretical developments since have largely followed the broader psychological zeitgeists of the times.

1.11.1 Early theoretical developments

Initial theories of test anxiety proposed by the pioneers in the field (Mandler & Sarason, 1952; Saraon & Mandler, 1952) were based on motivational or 'drive' theory (Hull, 1943). Drive theory is based on principle that we have psychological needs and we become tense or anxiously aroused when these needs are not met. Poorer examination performance was explained by excessive competition and interfering activation of competing responses when individuals were over-aroused. On the flip side, drive theory also states that high-test-anxious individuals will perform better than their peers on relatively easy tasks or test due to greater arousal (Spence & Spence, 1966). There is some support for the contention that high-test-anxious individuals perform worse their peers on difficult tasks (Eysenck, 1985; Spielberger, 1966). However, there is little evidence of a facilitative effect of anxiety on test performance (Tobias, 1985). The principal limitation of this model is that is difficult, if not impossible, to accurately assess drives and response thresholds and therefore it has low predictive utility (Zeidner, 1998).

1.11.2 Behavioural models of test anxiety

Application of behaviourist models boomed in the 1960 and 1970s. Classical conditioning, operant conditioning, and modelling theory underpin these models (Skinner, 1938; Watson, 1913; Bandura, 1965). Classical conditioning refers to learning whereby behaviour (e.g., anxiety) becomes paired with a stimulus (e.g., tests). Operant conditioning refers to behaviours that are modified by the effect they produce, such as reward (e.g., receiving praise) or punishment (e.g., failing a test). Modelling refers to how behaviours are learnt by observation of modelled or demonstrated behaviour. Poorer test performance is explained, like drive theory, by excessive and interfering response competition and that conditioned anxiety can lead to motivated attempts to avoid the stimulus to reduce anxiety. Support for behaviourist models largely comes from experimental and randomized controlled trials that show reductions in test anxiety after forms of behaviour modification (Ergene, 2003; Hembree, 1988). A key criticism of behaviourist models is that they fail to take account of cognitive processes that mediate between learning and behaviour (Zeidner & Matthews, 2011).

1.11.3 Study skills models

Study skills deficits models gained traction in the 1980s. They proposed that poorer test performance was attributable to poor or deficient study skills and not test anxiety (Benjamin et al., 1981; Culler & Holahan, 1980; Kirkland & Hollandsworth, 1980; Paulman & Kennelly, 1984). Test anxiety was considered an epiphenomenon borne out awareness was one was ill prepared to take the test,

and thus test anxiety is a correlate and not a casual factor in test performance. Although high-test-anxious students have been found to have poorer self-rated study skills than their peers (Culler & Holahan, 1980; Kirkland & Hollandsworth, 1979; Wittmaier, 1972), there is weak evidence that study skills competence is associated with test performance (Tobias, 1992). Similarly, there is inconsistent support for poorer test-taking strategies, with studies some high-test-anxious have poorer test-taking competence (e.g., Bruch, 1981), while other studies finding no differences (e.g., Bruch, 1983). Moreover, students with good study skill habits still report worry and cognitive interference in test situations, contradicting predictions of study skill models (Naveh-Benjamin et al., 1987). Moreover, test anxiety can influence study strategy, with greater test anxiety associated with increased surface-level learning (e.g., rote learning) (Spada & Moneta, 2012, 2014). Study skills models of test anxiety have fallen out of favour due to the disassociations between study skill ability and manifestation of test anxiety (Ziedner, 1998).

1.11.4 Cognitive and cognitive-behavioural models

Cognitive models proliferated following the identification of worry as the key feature of test anxiety (Doctor & Altman, 1969; Liebert & Morris, 1967). Cognitive models posit that thinking and other cognitive processes mediate the relationship between event (i.e., test) and response(s) (i.e., test anxiety). There are four main types of cognitive or cognitive-behavioural models: cognitive-attentional, transactional, schema, and self-regulation. Cognitive-attentional models of anxiety, and test anxiety, proposed that differences between high- and low-test-anxious students were due to the allocation and control of attention

(Wine, 1971; Sarason, 1980; Eysenck et al., 1992, 1997, 2007). Transactional process theory models of test anxiety (Spielberger & Vagg, 1995; Zeidner, 1998) conceptualizes stress as a dynamic process between person and environment, as well emphasizing the mediating effect cognitive appraisals have upon evaluating a situation as potentially threatening (i.e., primary appraisal) and then evaluating one's ability to cope with this threat (i.e., secondary appraisal) (Lazarus, 1991; Lazarus & Folkmam, 1984). Schema models of anxiety (e.g., Beck, Emery, & Greenberg, 1985) propose that individuals with anxiety have negative deep-seated or 'core beliefs' about the self, the world, and future (e.g., "I am a failure", "The world is a dangerous place", "Things will not get better") that influence behaviours. Finally, self-regulation models have also been applied to test anxiety (e.g., Carver & Scheier, 1991), which propose that maladaptive coping efforts maintain and exacerbate distress.

There is substantive support for the role of cognition in test anxiety and test performance (Cassady, 2002; Englert & Bertrams, 2015). High-test-anxious individual become self-preoccupied with their thoughts, and this interferes with test performance (Sarason, 1984, 1988). High-test-anxious students report greater negative self-focus, more test-irrelevant thinking, and more worry than their peers (Sarason, 1980; Sarason, Sarason, & Pierce, 1990; Wine, 1980). Experimental studies show that high-test-anxious students report greater worry and perform worse than their peers on tasks (Deffenbacher & Deitz, 1978; Derakshan, Smyth, & Eysenck, 2009), while observational studies using path analyses (e.g., Cassady, 2002; Heng-Tsung & Shao-Ting, 2013), and meta-analyses consistently find worry, but not emotionality, has a significant negative association with test performance (Hembree, 1988; Richardson et al., 2012;

Seipp, 1991; von der Embse et al., 2018). Key constructs implicated in the schema models, such as core beliefs, dysfunctional attitudes, and negative automatic thoughts have been implicated in test anxiety (Hunsley, 1985; Wong, 2008). Coping efforts moderate the effects of test anxiety on performance (Zeidner, 1996).

Overall, cognitive and information processing models are now the dominant explanatory paradigm in test anxiety.

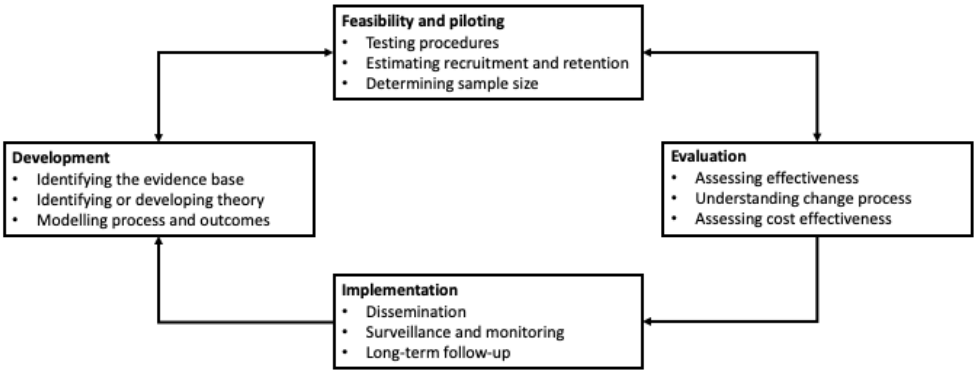
1.12 Summary

Test anxiety is a major educational concern in higher education. Tests determine the progression and prospects of students. Test anxiety undermines academic performance and is associated with poorer mental health. Understanding and developing effective interventions to treat test anxiety is an important priority. Despite test anxiety being a major educational concern there was a marked dip in research activity in the 1990s (Zeidner, 1998). However, due to increased testing in education settings because of government policy such as the 2001 “No child left behind” Government act in the USA, and increased demand on wellbeing services from students (Hubble & Bolton, 2020; Rückert, 2015), there has been a significant upturn in interest in test anxiety (von der Embse et al., 2018). Given limited resources available to education and wellbeing services, it is important to identify or develop effective interventions that treat test anxiety.

1.13 Thesis rationale and overarching research questions

Test anxiety has serious consequences for the educational prospects of university students and is associated with significantly poorer mental wellbeing. Therefore, it is vitally important that support and interventions are offered to students to help them reduce their test anxiety and allow them to fulfil their academic potential. The Medical Research Council (MRC; Craig et al., 2008; O'Cathain et al., 2019) offers a framework for the development and evaluation of complex interventions that is considered best practice, advocating a systematic approach to the development of interventions that is supported by best available evidence and grounded in theory. The framework consists of four phases (Figure 1.1).

Figure 1.1. Medical Research Council Framework for developing and evaluating complex interventions.



Briefly, Phase 1 comprises establishing what is known about current interventions and if new interventions are needed. If new interventions are needed, then theoretical models should be identified, applied, and empirically

tested. Phase 2 involves developing an intervention based upon the theoretical mode, examining its feasibility, and conducting pilot tests (e.g., to check on acceptability of the treatment and possible side effects). Phase 3 requires formal evaluation of the intervention, typically via a RCT. Finally, Phase 4 entails the implementation and evaluation of the intervention into routine practice.

The scope of this thesis largely covers Phases 1 and 2 of the MRC framework. The first phase involves establishing the efficacy of existing interventions, to determine if new interventions ought to be developed. Therefore, the first research question of this thesis is "how efficacious are existing interventions for reducing test anxiety and improving test and/or academic performance for test anxious university students?". This research question is addressed in Chapter 2.

The second phase of the MRC framework then specifies that new interventions that are developed should have a sound theoretical basis. Moreover, empirical support for the theory in the research context should be sought before investigations of an intervention efficacy are conducted. This leads to the second broad research question of "what models can account for the psychological mechanisms underpinning test anxiety?". This research question and the empirical examination of models that can account for the key features of test anxiety are addressed across Chapters 3-6. A summary of findings, conclusions, and implications of research are discussed in Chapter 7.

Chapter Two

Efficacy of interventions for test anxious university students: A meta-analysis of randomized controlled trials

2.1 Introduction

Given the negative impact test anxiety has on academic performance and student mental health it is not surprising that many interventions have been developed to treat it. Psychological interventions have been most commonly applied to treating test anxiety in undergraduate students (Zeidner, 1998). Test anxiety interferes with studying and test taking and therefore interventions that alleviate symptoms of test anxiety should improve test performance. Psychological interventions for test anxiety fall into two broad categories: behaviour therapy (BT) and cognitive-behavioural therapy (CBT).

Behavioural interventions are based on behaviourism, which emphasized outward and directly observable behaviour and rejected the inward experiential phenomena (Skinner, 1938; Watson, 1913). BT focuses on reducing the physiological arousal and somatic symptoms (i.e., emotionality dimension) of test anxiety. Interventions are therefore based on the assumption that behaviours (like feeling anxious in response to tests) are learnt, and that these behaviours can be changed by changing the consequences of behaviour. The goal is to reinforce desirable behaviours and reduce or eliminate undesirable ones. Additionally, new behaviours may need to be taught or modelled to people within interventions. There are four main types of behavioural strategies for altering unwanted behaviour: muscle relaxation, flooding, systematic desensitization, and modelling. In relaxation techniques such as progressive muscle relaxation (Jacobson, 1938) individuals are taught to sequentially relax certain muscle groups. Flooding involves exposing people to their fear-inducing object or situation and encouraging the person to endure their fear until it ameliorates. Flooding involves the immediate exposure to the feared stimulus or

situation (e.g., test, spider, confined space), with individuals needing to endure their fear reaction until it dissipates. Systematic desensitization involves counterconditioning techniques to teach muscle relaxation while visualizing a hierarchy of increasingly stressful test anxiety related scenes (Wolpe, 1958). Modelling involves people learning via observation of others' behaviour. Modelling can be explicit, where people are instructed to observation others, or implicit or covert where people observation others but are not informed of the nature of the experimentation.

The central idea in CBT is that distorted thoughts and disturbed thinking negatively influence mood and behaviour, and ultimately result in emotional issues and disorders. As thoughts, feelings, and behaviour are all connected, CBT approaches propose that identifying and changing unhelpful or faulty thinking will lead to improvements in behaviour and the emotional state of an individual. There are many varieties of CBT (David & Szentagotai, 2006; Hebert & Forman, 2011) but the most influential are rational emotive behavioural therapy (REBT; Ellis, 1957, 1962), and the schema-based approach of Beck and colleagues (CT; Beck, 1967; Beck et al., 1985). Both REBT and CT posit that it our thinking about events, and not the events themselves, that lead to emotional distress and disorders. The aim is to teach clients how to recognise and respond more adaptively to their thoughts. In the REBT, significant emotional issues are the result of irrational beliefs (e.g., self-defeating thinking such as "I cannot do this"). In CT, faulty or erroneous beliefs concern the self, the world, or the future (the cognitive triad), and can inferred by the negative automatic thoughts (NATs) that an individual reports in their thinking. The type of psychological disorder a person has is reflected in the content of their NATs, and this is referred to as the cognitive

specificity hypothesis (Baranoff & Oei, 2015; Beck et al., 1985). NATs are due to negative core beliefs stored in deeply held schemas. Schemas are memory structures organized around categories of information that can be updated with new information over time. However, CBT posits that the schemas about the self (e.g., “I am a failure”), world (e.g., “The world is a dangerous place”), and future (e.g., “Things will always be this way”) in psychological disorders have become more fixed and less changeable and that cognitive biases reinforce that worldview. Therapy for both REBT and CT focuses on challenging the validity and changing the content of these thoughts, termed ‘cognitive restructuring’. Commonly, behavioural experiments and techniques are included alongside cognitive restructuring, hence the term ‘CBT’.

There are other types of psychological therapy beyond BT and CBT. The most common are person-centred counselling (PCC; Rogers, 1961) and analytical psychotherapy (AP; Freud, 1900/2010). Unlike, BT and CBT, which focus upon the ‘here and now’ symptoms, both person-centred and psychoanalytical therapies propose that present distress and symptoms are due to deep-rooted issues that primarily related to childhood experiences. PCC is non-directive form of counselling with the client largely determining the scope and direction of the session content, whereas in AP the therapist will offer insight and guide the session. Both therapies consider the therapeutic relationship the key process for change (Ardito & Rabellino, 2011). Whilst BT and CBT-based therapies typically follow manuals or procedures over a set number of sessions, PCC and AP are ‘free form’, being guided by techniques and insights, and the number of sessions is not limited. It is perhaps for this latter reason that PCC and AP are seldom, if ever, applied to test anxiety; past meta-analyses did not report

any trials evaluating PCC or AP (Ergene, 2003; Hembree, 1988). Counselling interventions that have been evaluated in test anxiety have referred to group discussion sessions and have not explicitly referred to principles of PCC (e.g., Sapp, 1988).

Another class of interventions applied to treating test anxiety are study skill training (SST) interventions. These are based upon the assumption that inadequate test preparation and test taking skills impair academic performance and that test anxiety is an epiphenomenon arising from an individual's appraisal that they are inadequately prepared for the test. SST interventions therefore seek to enhance the learning and test-taking abilities to increase confidence and reduce anxiety in test taking situations so that individuals feel more confident about meeting their performance goals (Zeidner, 1998). SST interventions most commonly consist of two components (e.g., Dendato & Diener, 1986), one focusing on effective ways of learning and encoding study material (e.g., deeper-levels of learning as compared to rote-learning), with the other component focusing on effective strategies during examinations (e.g., allocating more time to those questions which represent a greater proportion of the total examination score).

Intervention packages have also been designed that combine psychological (BT, CBT) and SST interventions to address the psychological causes of test anxiety and improve effective study and test-taking skills (e.g., McCordick et al., 1981).

Finally, pharmacological approaches target the neurochemical origins of anxiety. Medications used to treat test anxiety include tranquilizers, beta-

blockers, anti-depressants, and benzodiazepines, though there is scant research on their application and efficacy (e.g., James & James 1973).

The efficacy of these interventions has been evaluated in randomized controlled trials (RCTs). RCTs are a research design in which enrolled participants are randomly assigned to intervention conditions. RCTs are considered the gold standard methodology for evaluating intervention efficacy as they permit causal inferences to be drawn.

2.1.1 Efficacy of interventions for test anxiety in university students

Two meta-analyses have examined the efficacy of interventions for test anxiety (Ergene, 2003; Hembree, 1988). Based on a review of 137 studies (n = 7,641 students) published between 1950 and 1986, Hembree (1988) concluded that BT and CBT were effective in reducing test anxiety and improving academic performance. Ergene (2003) examined the efficacy of interventions in 56 studies (n = 2,428 students) published between 1950 and 1998 focusing only on controlled trials, and concluded that BT, CBT, and Combined psychological and SST interventions were effective in reducing test anxiety (academic performance was not examined). Neither review included any pharmacological studies.

Despite the attempts of both meta-analyses to provide a systematic appraisal of the evidence at those time points, neither meta-analysis was conducted with the rigour expected of current systematic reviews, casting doubt on their conclusions. Both included and synthesized results from studies on primary/elementary school children through to university undergraduates. Clearly, symptom presentation will vary considerably across age ranges, and therefore the suitability and efficacy of interventions may also vary considerably.

Though Ergene (2003) reported a summary effect size of interventions for test-anxious undergraduates (*d* = -0.68, 95% CI -0.77 to -0.59), no summary statistics were reported by Hembree (1988), and neither review examined the efficacy of different treatment approaches for undergraduate students. Therefore, the efficacy of specific intervention approaches for specific student groups is unknown. It cannot be assumed that an intervention efficacious in one student group is equally efficacious in another group; it is possible efficacious interventions for one group may mask or moderate the effect for another group. Moreover, both reviews used a fixed-effect model, which assumes all their included studies are estimating the same effect, which cannot be the case given the age differences of samples across included trials (in addition to other methodological differences in trial conduct). Treatment trials of test anxiety have frequently included more than two treatments arms (e.g., Meichenbaum, 1972) but neither review reports how these trials are dealt with within their analyses, and therefore some participants, such as those in control conditions, may have been included more than once within their analyses. This biases results, by selectively increasing sample size for those studies with more than two conditions, giving them greater weight within data syntheses, and reducing confidence intervals around the effect size estimates (Borenstein et al., 2009). Finally, when viewed through the lens of modern reporting standards such as the Preferred Reporting Standards for Systematic Reviews and Meta-Analyses (PRISMA; Moher et al., 2009), the standard of both reviews is inadequate. Neither review identifies to the reader which trials were included and their basic characteristics (e.g., interventions delivered, delivery format, number of sessions), and neither uses forest plots to convey the variation in outcomes

between studies and which studies contributed most weight to the summary effect size. Additionally, *a priori* protocols are not available for either review.

Given these serious problems with the methodological and reporting standards of the two reviews above, the efficacy of interventions for test-anxious university students is unknown, and a new and up-to-date review that conforms to modern methodological and reporting standards is required.

2.1.2 Aims of meta-analytic review

This review has three main aims, to examine: (1) the efficacy of interventions for reducing test anxiety in university undergraduate students (henceforth 'students'), (2) potential moderators of treatment effect (see below for more detail) and, (3) the efficacy of interventions for improving academic performance.

2.1.2.1 Moderators

The following moderator variables would be examined provided sufficient data is available. Gender and pre-treatment severity will be examined as sample-dependent moderators of intervention effect. Given females report greater test anxiety than males (Cassady & Johnson, 2002; Zeidner, 1998), it was hypothesized that studies with higher proportions of female participants will have poorer treatment outcomes. Greater pre-treatment symptom severity is a predictor of poor treatment outcome in anxiety disorders (Knight et al., 2014). Therefore, it was also hypothesized that pre-treatment severity would be associated with poorer treatment outcomes in test anxiety.

Several methodological moderators of intervention effect are also examined: mode of treatment delivery (i.e., group vs. individual), treatment dosage (i.e., number of hours of treatment), and manualization of interventions (manualized vs. non-manualized). Psychological interventions are typically delivered face-to-face, either individually or in a group format. There are advantages and disadvantages to both delivery formats, with individual delivery affording greater privacy for discussing sensitive topics, whilst a group format permits the fostering of communication, social skills, and belonging among one other, as well being able to learn from others (DeLucia-Waack, Kalodner, & Riva, 2013). Group interventions were hypothesized to be associated with better treatment outcomes than individual interventions, in-line with Ergene (2003) findings that group interventions produced a larger effect size than individual interventions. Meta-analyses of psychological treatments find that duration of treatment has no significant relationship with treatment outcomes (Smith & Glass, 1977; Stiles, Barkham, & Wheeler, 2015). Similarly, Ergene (2003) found no dose-response pattern emerged between interventions for test anxiety and treatment outcomes, and therefore no significant relationship was hypothesized here also. Manualization involves the use of a structured treatment manual that specifies the content, principles, strategies, techniques, sequencing, and format of the intervention. As such manualization is intended to increase the replicability of treatment outcomes as independent therapists will practice using the same guidance (Kazdin, 2015). Manualization may improve the quality of treatment delivery (Addis et al., 2006), and therefore it was hypothesized that manualized interventions would be associated with better treatment outcomes.

2.2 Method

This review uses systematic review methods to search and screen for relevant RCTs, and meta-analysis statistical procedures to analyze and synthesize data. Systematic reviews use clear, *a priori* rules for searching and selecting studies, increasing transparency and reproducibility of results (Cook, Mulrow, & Haynes, 1997). Meta-analysis methods offer many advantages over the traditional analyses conducted in narrative reviews, in which a subject expert selects the studies for inclusion and conclusions would often be made on the basis on 'vote counting' methods (Freemantle & Geedes, 1998). In vote counting, the author would simply count and subjectively weigh up the balance of those studies with significant effects, and those without (Borenstein et al., 2009). Abelson wryly noted that narrative review data analyses are akin to "doing arithmetic with words" (cf. Borenstein et al., 2009, p.14). In fact, vote counting becomes more misleading the as the number of studies increases (Hedges & Olkin, 1980). Meta-analyses permit the computation of effect sizes for each study, facilitating the understanding of study results in the context of the other relevant studies, and if data are consistent, the computation of a summary effects. Summary effect sizes are a weighted mean of individual study effects. By synthesizing across studies, meta-analysis methods increase power to detect significant effects. Importantly, they also allow researchers to move beyond saying if a *p* value is significant or not, to draw conclusions about the size, direction, and clinical importance of effects, which speak to the practical importance of study findings.

The review was conducted and reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA; Moher

et al., 2009) guidelines. Development of the study protocol (Huntley et al., 2016) was guided by Cochrane protocol review guidelines (Cochrane, 2011) and was registered with PROSPERO (CRD42016035859).

2.2.1 Identification and selection of studies

2.2.1.1 Eligibility criteria

Included studies: (i) included only test-anxious university undergraduate students (self-selected and/or meeting criteria for severity of test anxious); (ii) examined psychological, SST, pharmacological interventions, and/or combined intervention packages; (iii) had at least one comparator that was a control condition (no treatment, waitlist, or psychological/pill placebo); (iv) reported a primary outcome of test anxiety severity on a valid measure; (v) used random assignment of participants to conditions; and (vi) was published in a peer-reviewed journal written in English.

2.2.1.2 Identification of studies

Six electronic databases – Cochrane Central Register of Controlled Trials (CENTRAL), Educational Resources Information Center (ERIC), MEDLINE, PsycINFO, Scopus, and Web of Science – were searched using variations of the relevant search terms covering: test anxiety; undergraduate student; psychological, SST, or pharmacological interventions; randomized controlled trial. An example search protocol for the Scopus database is presented below (Table 2.1) and the final search protocols can be viewed in Appendix A. The first search period covered all years up to April 2016. A further search was conducted, covering the period April 2016 up to May 2017.

A manual search was conducted, checking reference lists of previous meta-analytic reviews and primary articles. Articles identified through the literature searches were then combined into a single database and duplicates removed.

2.2.1.3 Screening of studies

Two researchers, Dr James Temple (JT) and Dr Melissa Longworth (ML) independently screened studies against eligibility criteria, with discrepancies resolved by the thesis author. First, studies were screened by title and abstract. Full-text copies of the remaining included studies were then retrieved and screened using tool with items for each of the eligibility criteria. Studies could be then marked as 'Included', 'Maybe', or 'Excluded'. Studies marked as maybe were reviewed by the thesis author.

2.2.2 Data extraction and risk of bias assessment

Predefined information was independently extracted by JT and ML from the study reports using a modified Cochrane data extraction form. Risk of bias (RoB) was assessed by Cochrane's RoB tool (Higgins, 2011) that assesses the following domains: random sequence generation, allocation concealment, blinding of participants and personnel, blinding of outcome assessment, incomplete outcome data, and selective reporting.

Each domain is scored either: 'low risk of bias', 'unclear', or 'high risk of bias'. If relevant information was not reported, study authors were contacted by the author and the missing study information requested. If there was no response

Table 2.1. Example search protocol (for Scopus database).

#	Search Terms
1	"Test anxiety" OR "Examination anxiety" OR "Test stress" OR "Examination stress" OR "Test performance" OR "Examination performance" OR "Evaluation Stress" OR "Evaluative Stress" OR "Evaluative Anxiety" OR "Evaluation Anxiety" OR "Test fear" OR "Examination fear"
2	College student OR University student OR undergraduate
3	Randomi?ed controlled trial OR Controlled clinical trial OR Clinical trial OR PHASE* Trial OR RCT OR randomi?ed OR random allocation OR randomly allocated OR Controlled trial OR placebo OR outcome study
4	Psycho* or behavi* or cogniti* or metacogniti* or implosive or desensiti* or interpersonal or gestalt or attenti* or activation* or rational or bibliotherapy or counsel* or supportive or study skill* or skill focus* or testwise* or relax* or training* or nondirective or guided or imag* or computer* or CBT or pharmaco* or drug* or treatment or intervention
5	#1 AND #2 AND #3 AND #4
6	Limit #5: English Language

within two weeks, a reminder email was sent. Following this, no further efforts were made to obtain information.

2.2.3 Planned Analyses

Meta-analyses were conducted using Revman 5.3.5 (Cochrane Collaboration, 2014), while tests of publication bias and meta-regression

analyses were conducted in R 3.3.3 (R Core Team, 2013) using the 'metafor' package (Viechtbauer, 2010).

2.2.3.1 Data synthesis

A random-effects model was used. Test anxiety severity is measured on a continuous scale, but a wide variety of self-report outcome measures have been used across studies. Therefore, between-study standardized mean differences (SMD) with 95% confidence intervals were calculated for each pairwise comparison (active intervention versus control group) and for the pooled effect, using Cohen's effect size with Hedges' correction for small sample bias i.e., Hedges' *g* (Hedges, 1981).

Studies with more than two treatment comparisons were included. However, independence between multiple comparisons within the same study cannot be assumed (Borenstein et al., 2009). To adjust for the potential bias the control group *n* is divided by the number of active comparator conditions (Cochrane, 2011) ensuring each participant is only included once in the analyses. Data from the primary test anxiety outcome measure reported in each study were used (if no primary outcome stated, the first reported test anxiety outcome was then used).

The magnitude of heterogeneity was assessed using the I^2 statistic, with values greater than 50% indicative of at least moderate heterogeneity (Higgins et al., 2003). Subgroups analyses were used to examine the relative efficacy of different interventions at post-treatment and follow-up. Using past reviews as a guide, the treatment categories were decided as the following: BT, CBT, SST,

and Combined. Combined interventions consisted of both psychological (i.e., BT or CBT) and SST components.

The same procedure outlined above was conducted for the academic performance analyses. Academic performance indices used here included either post-treatment examination results or grade point average (GPA). GPA is summary statistic representing the average value of accumulated grades earned across course modules. Most commonly, GPA is based on a 0 to 4.0 scale (A = 4.0, B = 3.0, C = 2.0, D = 1.0, Fail = 0.0) where a 4.0 GPA average represents the highest score.

Outliers were identified via visual inspection of the forest plot (i.e., where study effect estimates, and their 95% confidence interval, had little overlap with other study effect size estimates). Outliers were then removed, and results compared with original estimates.

2.2.3.2 Moderator analyses

Moderators under investigation have been discussed previously (see section 2.1.2.1). A minimum of 10 studies for each study-level variable is recommended (Higgins & Green, 2011). Therefore, if heterogeneity was present within any of the subgroups, and there were sufficient studies, moderators of treatment outcome were examined via mixed-effects meta-regression (i.e., studies are treated as being a random sample, but the moderators are treated as being fixed-effect). Regarding manualization, a study had to explicitly cite a manual or provide a detailed description within the manuscript to be coded as "manualized"; else they were coded as non-manualized.

2.2.3.3 Sensitivity analyses

Sensitivity analyses were planned to assess the robustness of results, with RCTs considered at high risk of bias removed and then results compared to original analyses.

2.2.3.4 Assessment of publication bias

Publication bias was assessed using Egger's test (Egger et al., 1997) and funnel plots of the effect sizes against the standard error (SE).

2.2.4 Quality of evidence

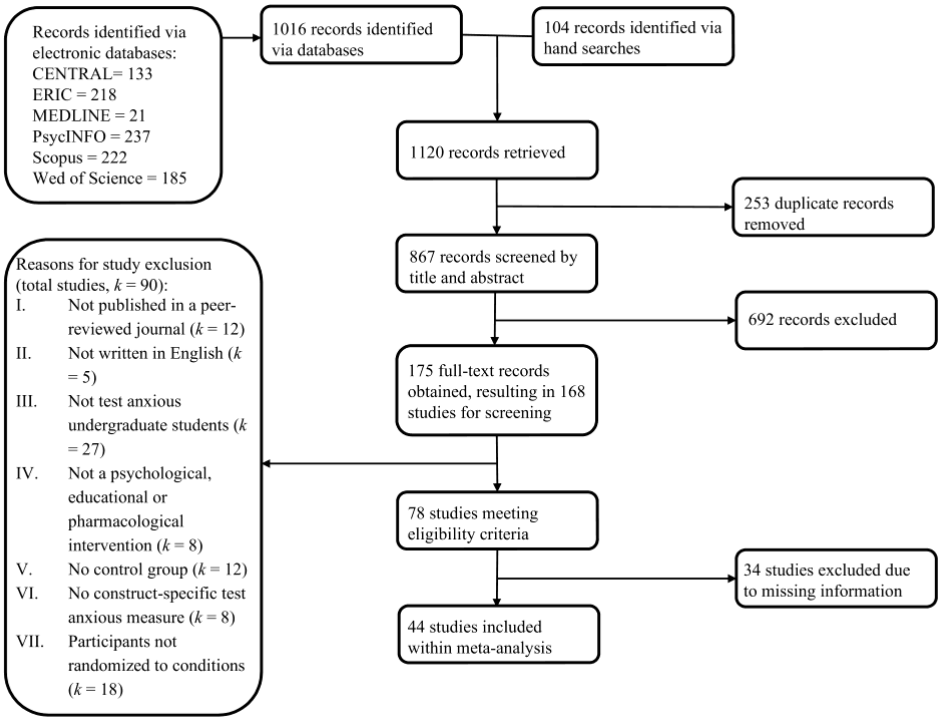
Overall quality of evidence was assessed with Grading of Recommendations Assessment, Development and Evaluation (GRADE; Guyatt et al., 2008) using their GRADEpro Guideline Development Tool (GDT). Five domains of quality are assessed: risk of bias, inconsistency of trial results, indirectness of measurement, imprecision of effect size estimates, and the likely impact of publication bias. Evidence quality is graded on a 4-point scale from 'very low quality' (i.e., limited confidence in effect size estimate) to 'high quality' (i.e., high confidence in the effect size estimate).

2.3. Results

2.3.1 Identification and selection of studies

Figure 2.1 shows the PRISMA flowchart. The search identified 1,120 papers.

Figure 2.1. PRSIMA flowchart.



After removing 253 duplicate records, 867 records were independently screened (JT, ML) by title and abstract, with 692 records excluded. Full-text copies of the remaining 175 papers were obtained, with multiple papers corresponding to a single study grouped together, resulting in 168 unique studies. Full-text articles were independently screened (JT, ML), with 78 studies meeting the inclusion criteria; inter-rater reliability was good (Kappa = 0.60). The list of studies excluded (including reasons for exclusion) at the full-text screening stage can be found in Appendix B. The 78 studies were then examined to check if data necessary for inclusion in the meta-analyses was reported. Thirty-four studies did not report sufficient information; the study authors were contacted and asked to supply the required information. Although 16 (47%) of the study authors responded, none could provide the necessary information; primarily because

their studies were conducted more than 20 years ago, and data was no longer accessible. Forty-four studies were therefore included in the meta-analysis.

2.3.2 Characteristics of included studies

Descriptive information about the studies included within the meta-analyses is detailed in Table 2.2.

Most studies, 75%, were conducted in the United States of America (*k* studies = 33), with 25% conducted elsewhere (*k* = 11). Thirty-eight studies were published between 1970 and 1999, with only six studies published between 2000 and 2017.

There were 152 intervention arms across the 44 studies, of which 91 were active intervention conditions and 61 were control conditions. BTs were most frequently examined (across 59 conditions, *k* = 33), followed by CBTs (14 conditions, *k* = 11), SST (10 conditions, *k* = 10), and Combined interventions (8 conditions, *k* = 4). Of the control conditions, no treatment control (NTC) was most frequently used (across 29 conditions, *k* = 29), followed by active control conditions designed to resemble placebo conditions (17 conditions, *k* = 17), and waitlist control (WLC; 15 conditions, *k* = 15). Active control conditions included instructional control, non-directive supportive counselling, and interventions with the purportedly 'active' ingredients removed such as implosive treatment based upon neutral cues rather than test anxiety specific cues. No pharmacological interventions were evaluated in any of the included studies (and only one study was identified in the searches but which did not meet the eligibility criteria).

Group delivery was used most often (76 conditions, *k* = 34), followed by individual interventions (14 conditions, *k* = 10), and computer-based delivery (one condition, *k* = 1). The overall duration of the interventions varied considerably, ranging from one hour to 26 hours (*M* = 7.24 hours, *SD* = 4.78). Although most studies (*k* = 30, 68%) did report using cut-off points for test anxiety, this varied greatly from study-to-study, with studies enrolling participants based on: (i) scoring above a cut-off score (*k* = 19), (ii) scoring above a set percentile for test anxiety severity (ranging from 50th to 85th percentile across studies) (*k* = 9), (iii) scoring within a range of scores for test anxiety severity (*k* = 1), or (iv) meeting diagnostic criteria for DSM-IV (American Psychiatric Association, 1994) specific phobia or social phobia (*k* = 1). Participants self-selected into studies that did not use cut-off criteria (*k* = 14) by responding to announcements or advertisements regarding test anxiety trials.

Reporting of the flow of participants within trials was poor; attrition was only reported in 27 of the 44 studies (2, 5-6, 11-15, 17, 19-23, 25, 27, 30-34, 36, 38-41, 44). The mean number of participants enrolled into active treatments was 15.8 (*SD* = 9.8, range 7 to 48), and for control conditions was 17.1 (*SD* = 11.0, range 7 to 48). Mean attrition for active conditions was 11.3% (*SD* = 12.0, range from 0% to 45%), and for control conditions was 7.8% (*SD* = 10.3, range from 0% to 35%). Only 15 of the 44 (34%) studies reported follow-up data (1, 4, 6-7, 18-19, 21-22, 24-25, 29, 32, 38-39, 26), with a mean follow-up of 9.4 weeks (*SD* = 6.4, range from 3 to 26 weeks). Three studies (21, 24-25) had a second follow-up period (*M* = 56 weeks, *SD* = 6.9). Only two studies (41-42) included a Consolidated Standards of Reporting Trials (CONSORT; Schulz et al., 2010) flow diagram.

Table 2.2. Selected characteristics of included studies.

ID	First Author - Year Country	Intervention(s)	Delivery format	Sessions	TA criteria used to select participants	Primary outcome measure	Follow-up period
#1	Mitchell 1970 AUS	Systematic Desensitization	Group	10 x 1 h (twice weekly)	≥ 1 SD above M on AAT-D	AAT-D	14 weeks
		Waitlist Control	-	-			
		No Treatment Control	-	-			
#2	Lomont 1971 USA	Systematic Desensitization	Group	8 x 1 h (weekly)	> 30 on AAT-D & < 23 on AAT-F	AAT-D	-
		Supportive Counselling	Group	8 x 1 h (weekly)			
		No Treatment Control	-	-			
#3	Prochaska 1971 USA	Implosive Therapy – TA Symptoms	Group	3 x 1 h (fortnightly)	None	AAT-D	-
		Implosive Therapy – TA Dynamic	Group	3 x 1 h (fortnightly)			
		Implosive Therapy – General Anxiety	Group	3 x 1 h (fortnightly)			
		Instructional Control	Group	3 x 1 h (fortnightly)			
		No Treatment Control	-	-			
#4	Meichenbaum 1972 CAN	Cognitive modification	Group	8 x 1 h (weekly)	None	AAT-D	4 weeks
		Systematic Desensitization	Group	8 x 1 h (weekly)			

		Waitlist Control	-	-			
#5	Dawley 1973	Implosive Therapy	Group	5 x 0.5 h (twice weekly)	≥ 212 on TAQ		-
	USA	Instructional Control	Group	5 x 0.5 h (twice weekly)	(66th percentile)		
		No Treatment Control	-	-			
#6	Kostka 1974	Systematic Desensitization	Group	10 x 1 h (twice weekly)	NR	STABS	20 weeks
	USA	Covert Modelling	Group.	8 x 1 h (twice weekly)			
		No Treatment Control	-	-			
#7	Guidry 1975	Covert Reinforcement	Group	1 h + 5 x 0.5 h	None	STABS	3 weeks
	USA	Instructional Control	Group	1 h + 5 x 0.5 h			
		No Treatment Control	-				
#8	Mitchell 1975	Systematic Desensitization + Study Skills Training	Group	26 x 1 h (twice weekly)	None	AAT-D	-
	AUS	Relaxation Therapy + Study Skills Training	Group	26 x 1 h (twice weekly)			
		Study Skills Training	Group	26 x 1 h (twice weekly)			
		Instructional Control	Group	5 x 1 h (twice weekly)			
		No Treatment Control	-	-			
#9	Anton 1976	Systematic Desensitization	Group	8 x 1 h (weekly)	None	STAI-T	-
	USA	Supportive Counselling	Group	4 x 1 h (weekly)			

		No Treatment Control	-	-			
#10	Bedell 1976	Systematic Desensitization -High	Group	7 x 1 h (weekly)	NR	TAS	-
	USA	Systematic Desensitization - Low	Group	7 x 1 h (weekly)			
		Relaxation Therapy -High	Group	7 x 1 h (weekly)			
		Relaxation Therapy -Low	Group	7 x 1 h (weekly)			
		No Treatment Control	-				
#11	Chang-Liang 1976	Systematic Desensitization	Group	3 x 1 h (weekly)	≥ 66th	STABS	-
	USA	Applied Relaxation	Group	3 x 1 h (weekly)	percentile		
		Relaxation Therapy	Group	3 x 1 h (weekly)			
		No Treatment Control	-	-			
#12	Melnick 1976	Systematic Desensitization	Group	4 x 1 h (weekly)	≥ 50th	TAQ	-
	USA	Attentional Control	Group	4 x 1 h (weekly)	percentile on		
		No Treatment Control	-	-	TAQ		
#13	Russell 1976	Systematic Desensitization	Group	5 x 1 h (weekly)	None	TAS	-
	USA	Cue Controlled Relaxation	Group	5 x 0.5 h (weekly)			
		No Treatment Control	-	-			
#14	Denney 1977	Self-Control Desensitization – Active	Ind.	8 x 1 h (weekly)	≥ 85th	STABS	-

	USA	Self-Control Desensitization – Passive	Ind.	8 x 1 h (weekly)	percentile on		
		Systematic Desensitization – Active	Ind.	10 x 1 h (weekly)	TAQ		
		Systematic Desensitization – Passive	Ind.	10 x 1 h (weekly)			
		Instructional Control	Ind.	8 x 1 h (weekly)			
		No Treatment Control	-	-			
#15	Finger 1977	Attention Training	Group	8 x 1 h (twice weekly)	≥ 66th	AAT-D	-
	USA	Relaxation Therapy	Group	8 x 1 h (twice weekly)	percentile on		
		Attention Training + Relaxation Therapy	Group	8 x 1 h (twice weekly)	AAT-D –		
		Waitlist Control	-	-	AAT-F scores		
#16	Horne 1977	Systematic Desensitization	Group	10 x 1 h (twice weekly)	100 highest	TAS	-
	USA	Flooding	Group	10 x 1 h (twice weekly)	scorers on		
		Modelling	Group	10 x 1 h (twice weekly)	TAS (out of		
		Study Skills Training	Group	10 x 1 h (twice weekly)	175 screened)		
		No Treatment Control	-	-			
#17	Counts 1978	Electromyographic + Cue Controlled Relaxation	Ind.	6 x 1 h (thrice weekly)	≥ 21 on TAS	TAS	-
	USA	Cue Controlled Relaxation	Ind.	6 x 1 h (thrice weekly)			
		Attentional Control	Ind.	6 x 1 h (thrice weekly)			

		No Treatment Control	-	-			
#18	Gallagher 1978	Covert Modelling	Ind.	4 x 0.5 h (twice weekly)	≥ 21 on TAS	TAS	3 weeks
	USA	Imagined Modelling	Ind.	4 x 0.5 h (twice weekly)			
		Waitlist Control	-	-			
#19	Lent 1978	Study Skills Training + Systematic Desensitization	Group	5 x 1 h (weekly)	Top 30% on	TAS	6 weeks
	USA	Study Skills Training + Cue Controlled Desensitization	Group	5 x 1 h (weekly)	TAS & SSHA		
		Study Skills Training	Group	5 x 1 h (weekly)			
		No Treatment Control	-	-			
#20	Romano 1978	Electromyographic feedback + Systematic Desensitization	Group	13 x 0.5 h (twice weekly)	≥ 70th	STABS	-
	USA	Electromyographic feedback	Group	13 x 0.5 h (twice weekly)	percentile		
		Systematic Desensitization	Group	13 x 0.5 h (twice weekly)	on STABS		
		Waitlist Control	-	-			
#21	Deffenbacher 1979a	Systematic Desensitization	Group	7 x 1 h (weekly)	> 34 on	AAT-D	6 weeks
	USA	Self-Control Relaxation	Group	7 x 1 h (weekly)	AAT-D		52 weeks
		Waitlist Control	-	-			
		No Treatment Control	-	-			
#22	Deffenbacher 1979b	Systematic Desensitization	Group	8 x 1 h (twice weekly)	> 25 on	TAS	9 weeks

	USA	Self-Control Desensitization	Group	8 x 1 h (twice weekly)	AAT-D		
		Waitlist Control	-	-			
#23	Holahan 1979	Cognitive Therapy	Group	4 x 2 h (weekly)	Top 20% on	TAS	-
	USA	Anxiety Management Training	Group	4 x 2 h (weekly)	TAS		
		Waitlist Control	-	-			
#24	Deffenbacher 1980a	Anxiety Management Training	Group	6 x 1 h (weekly)	Top 10% on	AAT-D	6 weeks
	USA	No Treatment Control	-	-	AAT-D		52 weeks
#25	Deffenbacher 1980b	Anxiety Management Training	Group	6 x 1 h (weekly)	> 37 on	AAT-D	6 weeks
	USA	Self-Control Desensitization	Group	6 x 1 h (weekly)	AAT-D		64 weeks
		Waitlist Control	-		(Top 5%)		
		No Treatment Control	-	-			
#26	Levine 1980	Negative Practice	Group	6 x 1 h (weekly)	> mean on	TAQ	-
	USA	Negative Practice + Homework	Group	6 x 1 h (weekly)	TAQ		
		Systematic Desensitization	Group	6 x 1 h (weekly)			
		Attentional Control	Group	6 x 1 h (weekly)			
		Waitlist Control	-	-			
#27	Lurie 1980	Covert Reinforcement	Group	6 x 1 h (twice weekly)	≥ 18 on TAS	TAS	-

	USA	Covert Rehersal	Group	6 x 1 h (twice weekly)			
		Study Skills Training	Group	6 x 1 h (twice weekly)			
		Attentional Control	Group	6 x 1 h (twice weekly)			
		Waitlist Control	-	-			
#28	Reed 1980	Biofeedback	Ind.	8 x 0.5 h (twice weekly)	≥ 31 on	AAT-D	-
	USA	Instructional Control	Ind.	8 x 0.5 h (twice weekly)	AAT-D		
		No Treatment Control	-	-			
#29	Altmaier 1981	Systematic Desensitization	Group	6 x 1 h (twice weekly)	≥ 50 on	TAI	12 weeks
	USA	Study Skills Training	Group	6 x 1 h (twice weekly)	TAI		
		Systematic Desensitization + Study Skills Training	Group	12 x 1 h (twice weekly)			
		No Treatment Control	-	-			
#30	Barabasz 1981	Rational Emotive Therapy	Group	4 x 1 h (twice weekly)	High scorers	TAQ	-
	NZ	Study Skills Training	Group	4 x 1 h (twice weekly)	on skin		
		No Treatment Control	-	-	conductance		
#31	D'Alelio 1981	Cognitive Therapy – 8	Group	8 x 1.5 h (weekly)	≥ 55 on TAI	TAI	-
	USA	Cognitive Therapy – 4	Group	4 x 1.5 h (weekly)			
		No Treatment Control	-	-			

#32	Decker 1981	Cognitive-Behavioural Therapy	Group	4 x 1.5 h (weekly)	≤ mean of	TAS	10 weeks
	USA	Study Skills Training	Group	4 x 1.5 h (weekly)	TAS		
		Waitlist Control	-	-			
#33	McCordick 1981	Cognitive-Behavioural Therapy + Study Skills Training	Group	15 x 1 h (twice weekly)	None	AAT-D	-
	USA	Cognitive Therapy + Study Skills Training	Group	15 x 1 h (twice weekly)			
		Systematic Desensitization + Study Skills Training	Group	15 x 1 h (twice weekly)			
		Study Skills Training	Group	1 x 1 h (weekly)			
		Waitlist Control	-	-			
#34	Russell 1982	Systematic Desensitization	Group	5 x 1 h (weekly)	None	TAS	-
	USA	Cue Controlled Relaxation	Group	5 x 1 h (weekly)			
		Systematic Desensitization + Cue Controlled Relaxation	Group	5 x 1 h (weekly)			
		Attentional Control	Group	5 x 1 h (weekly)			
		No Treatment Control	-	-			
#35	Ricketts 1984	Progressive Muscle Relaxation	Group	1 x 1 h	≥ 147 on	STABS	-
	USA	Study Skills Training	Group	1 x 1 h	STABS		
		Rational Emotive Therapy	Group	1 x 1 h			
		Instructional Control	Group	1 x 1 h			

#36	Crowley 1986	Cognitive Therapy	Group	1 x 6 h	≥ 32 on AAT	TAI	-
	USA	Cognitive Therapy	Ind.	6 x 1 h (twice weekly)			
		No Treatment Control	-	-			
#37	Bauman 1994	Eye Movement Desensitization Reprogramming	Ind.	1 x 1 h	TAI ≥ 50th	TAI	-
	CAN	No Treatment Control	Ind.	1 x 1 h	percentile		
#38	Sapp 1996	Relaxation Therapy	Ind.	8 x 1 h (weekly)	None	TAI-W	8 weeks
	USA	Supportive Counselling	Ind.	8 x 1 h (weekly)			
#39	Maxfield 2000	Eye Movement Desensitization Reprogramming	Ind.	1 x 1.5 h	≥ 50 on TAI	TAI	8 weeks
	CAN	Waitlist Control	-	-			
#40	Orbach 2007	Cognitive-Behavioural Therapy	Comp.	6 x 0.5 h (weekly)	None	TAI	-
	UK	No Treatment Control	-	-			
#41	Raijah 2014	Progressive Muscle Relaxation + Systematic Desensitization	Ind.	6 x 1 h (twice weekly)	None	WTAS	-
	MAL	No Treatment Control	Ind.	-			
#42	Saravanan 2014	Progressive Muscle Relaxation + Systematic Desensitization	Ind.	5 x 1 h (twice weekly)	Score of 30-39	WTAS	-
	MAL	Waitlist Control	-	-	on WTAS		
#43	Hahm 2016	Study Skills Training	Group	14 x 1 h (weekly)	None	G-TAI	-
	GER	No Treatment Control	-	-			

#44	Reiss 2017	Cognitive-Behavioural Therapy	Group	5 x 3 h (weekly)	DSM-IV	G-TAI	26 weeks
	GER	Cognitive-Behavioural Therapy + Imagery Re-scripting	Group	5 x 3 h (weekly)	SCID-I (social		
		Self-help	Group	5 x 3 h (weekly)	/specific phobia)		

Notes.

AAT-D/F = Achievement Anxiety Test – Debilitating/Facilitating; Comp = Computer/internet delivered; DSM-IV = Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition; G-TAI = German language version of TAI; h = hour(s); NR = Not Reported; RTAS = Revised Test Anxiety Scale; SCID-I = Structured Clinical Interview DSM-IV Axis I Disorders; SSHA = Survey of Study Habits and Attitudes; STABS = Suinn Test Anxiety Behaviour Scale; TA = Test Anxiety; TAI = Test Anxiety Inventory; TAI-W = TAI – Worry (subscale); TAQ = Test Anxiety Questionnaire; TAS = Test Anxiety Scale; WTAS = Westside Test Anxiety Scale.

Ten different test anxiety measures were used across the studies, with the Achievement Anxiety Test – Debilitating (AAT-D; Alpert & Haber, 1960) administered most frequently (*k* = 16). A single test anxiety measure was administered in 29 studies, with two or more test anxiety measures administered in 15 studies. There was variation in pre-intervention scores in measures across studies (e.g., between 30.74 and 41.06 on AAT-D). However, given the wide range of outcome measures, the lack of normative data, and no diagnostic criteria for test anxiety, it is difficult to interpret severity of pre-treatment test-anxiety. Four studies administered a state measure of anxiety in an examination context. The effect of interventions on improving academic performance was examined across 17 studies, via students' GPA scores (*k* = 11, studies 3, 9, 16, 21, 23-26, 32, 38, 41), examination scores shortly after treatment (*k* = 4, studies 2, 8, 12, 37), or both (*k* = 2, studies 14, 33).

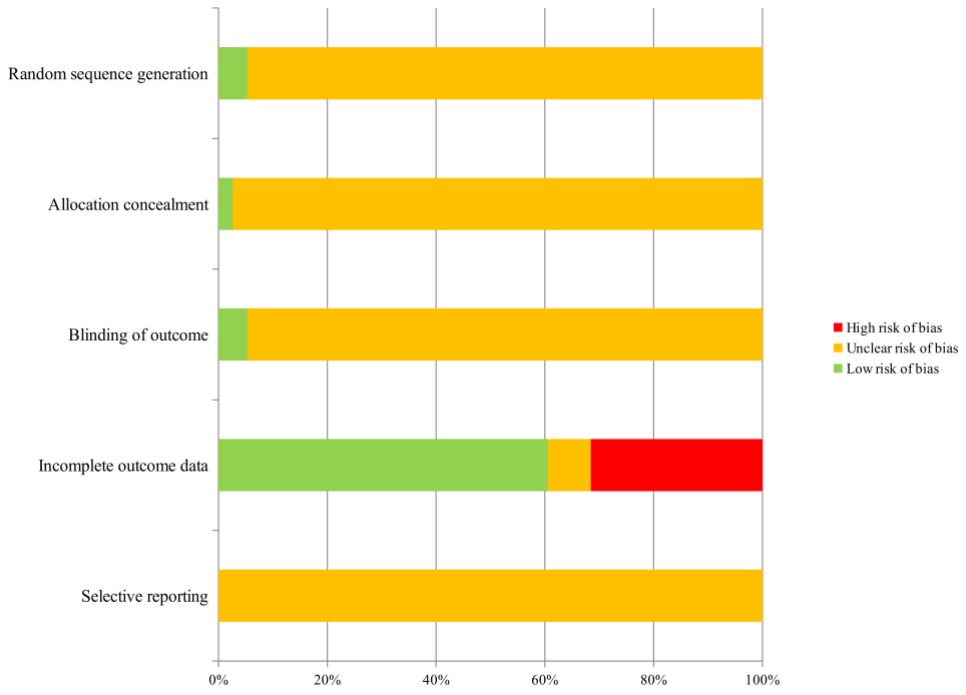
Only 27 studies reported the number of male and female participants that were enrolled. Gender proportion varied considerably, with percentages of females ranging from 0% to 97% (*M* = 64.22, *SD* = 19.28). Only one study (44) reported the proportion of male and female participants for specific intervention conditions. Just six studies (14%, studies 8, 17, 19, 40, 42, 44) reported the mean age of the sample, and only two (5%, studies 40, 44) reported mean age for specific intervention conditions.

Finally, manuscripts were examined to see: (i) if any adverse events had been reported, and (ii) if the clinical significance of the results had been evaluated. No adverse events were reported, nor was the clinical significance of study results evaluated.

2.3.3 Risk of bias assessment

Figure 2.2 summarizes the risk of bias assessments across all included studies.

Figure 2.2. Summary of risk of bias assessments.



As all included trials used psychological and educational interventions, the blinding of participants items was not included. Overall, 79% of items were scored as 'unclear' indicating poor quality of reporting. Only two studies reported how the random sequence was generated (4.5%), and only one study reported use of allocation concealment (2.3%). Five studies used independent personnel to assess study outcomes (11.4%). Twenty-seven studies (61.4%) were considered to have a low risk of bias with regard to incomplete outcome data. No study referred to a trial protocol. Inter-rater reliability between the independent

raters (JT and ML) was good (Kappa = 0.67). Complete risk of bias assessments for each study can be found in Appendix C.

2.3.4 Meta-analyses

2.3.4.1. Categorisation of interventions

Interventions were categorised as follows: (i) BT: anxiety management training, applied relaxation, biofeedback, covert modelling, covert rehearsal, covert reinforcement, cue controlled relaxation, electromyographic-based interventions, eye movement desensitization reprogramming, flooding, imagined modelling, implosive therapy, negative practice, progressive muscle relaxation, relaxation therapy, modelling, self-control desensitization, and systematic desensitization; (ii) CBT: cognitive therapy, cognitive-behavioural therapy (including with imagery rescripting), attention training, and rational emotive therapy; and (iii) Control conditions: attentional control, instructional control, no treatment control, self-help, supportive counselling, and waitlist control. There were no distinct variations of SST. Combined interventions comprised a form of BT or CBT plus SST.

2.3.4.2 Efficacy of interventions for reducing test anxiety, at post-treatment

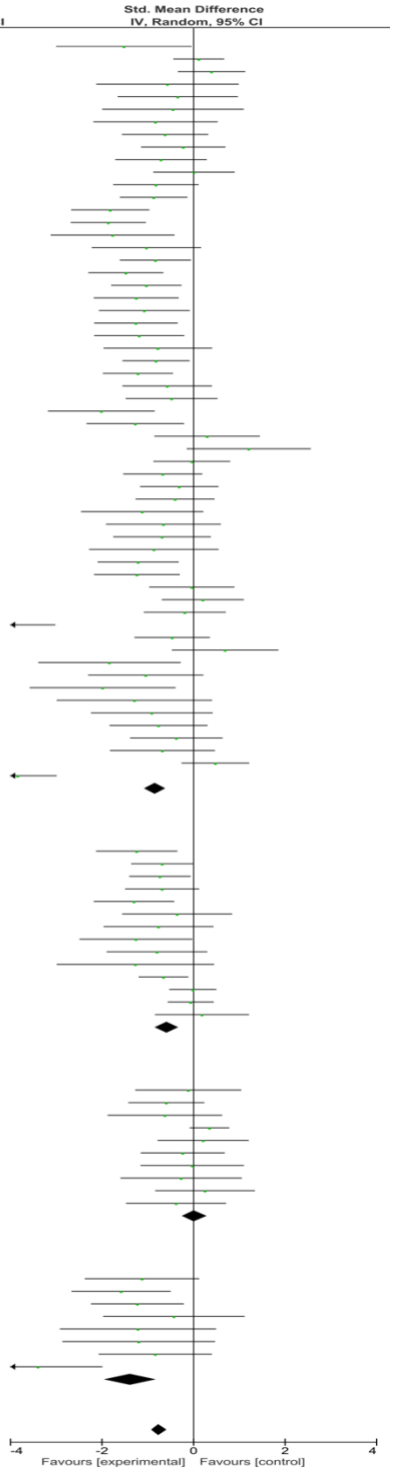
From the 44 studies included, 91 active-to-control group comparisons were included within the analyses of post-treatment data. A total of 2,209 participants were included in the analyses, with 1,362 participants in active treatment conditions and 847 participants in control conditions. Overall, interventions for test anxiety were superior to control conditions in reducing test anxiety severity (see Figure 3; see also Appendix D, Figure D.1 for annotated

forest plot, and Figure D.2 for the funnel plot), with an overall standardized mean effect (Hedges' *g*) across all interventions of -0.76 (95% CI -0.93 to -0.58). As expected, significant homogeneity was indicated (I^2 = 66%), and therefore a subgroups analysis was conducted. Combined interventions (*g* = -1.38, 95% CI -1.96 to -0.81, $p < .001$, *k* = 8), BT (*g* = -0.83, 95% CIs -1.06 to -0.60, *k* = 59, $p < .001$), and CBT (*g* = -0.58, 95% CIs -0.83 to -0.33, $p < .001$, *k* = 9) were significantly superior to control conditions. SST delivered alone (*g* = 0.02, 95% CIs -0.25 to 0.28, *p* = .900, *k* =10) was not significantly different from control conditions.

Inspection of the forest plot identified three outliers: two in the BT subgroup (studies 41-42) and one in the Combined approach (study 8; SD+SST intervention). Removal of these studies resulted in the overall effect estimate decreasing to *g* = -0.64 (95% CIs -0.77 to -0.50), the BT subgroup effect size changing to *g* = -0.70 (95% CIs -0.88 to -0.52), and the Combined subgroup estimate changing to *g* = -1.14 (95% CIs -1.62 to -0.66).

Figure 2.3. Forest plot of efficacy of interventions versus control conditions for reducing test anxiety at post-treatment

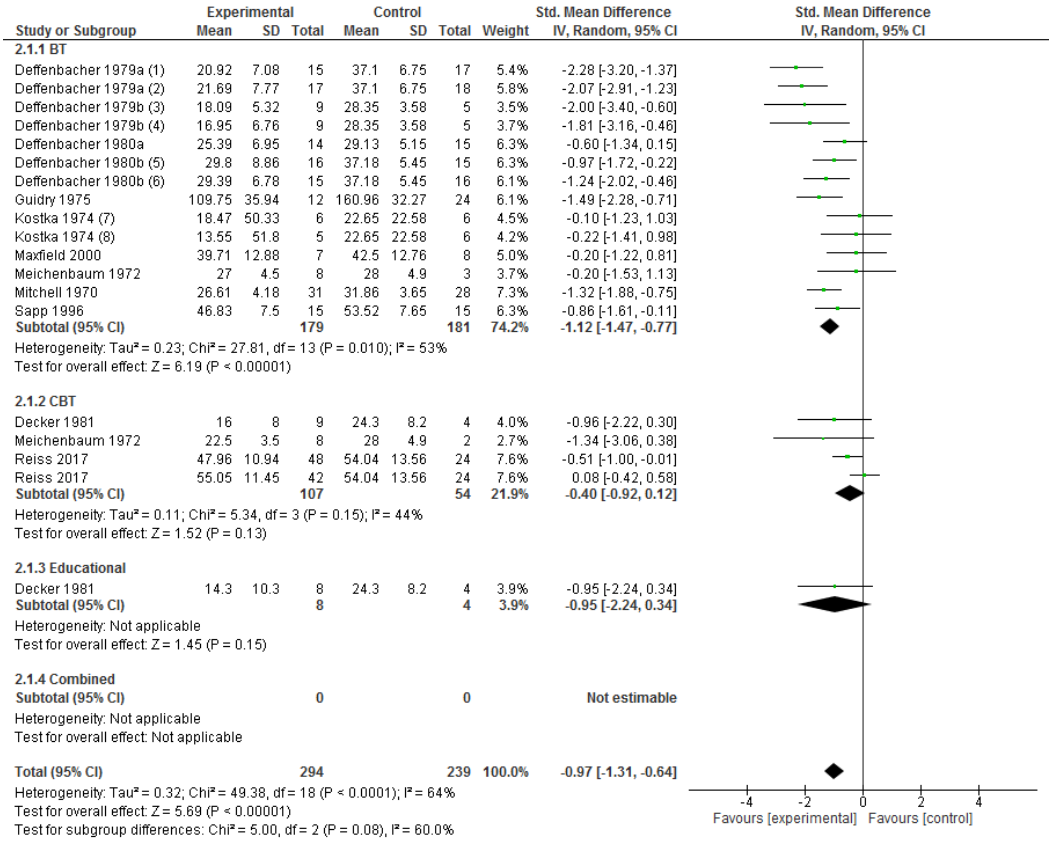
Study or Subgroup	Experimental Mean	Experimental SD	Experimental Total	Control Mean	Control SD	Control Total	Weight	Std. Mean Difference IV, Random, 95% CI
1.1.1 BT								
Altmaier 1981	41.1	7.92	10	54	7.73	3	0.8%	-1.52 [-2.99, -0.05]
Anton 1976	45.5	8.84	32	44.36	11.26	22	1.5%	0.11 [-0.43, 0.66]
Bauman 1994	51.33	13.23	15	47	7.41	15	1.4%	0.39 [-0.33, 1.12]
Bedell, 1976	22.9	5.04	10	26.1	6.33	2	0.7%	-0.57 [-2.11, 0.97]
Bedell, 1976	23.5	7.15	10	26.1	6.33	3	0.9%	-0.35 [-1.65, 0.95]
Bedell, 1976	21.9	8.84	10	26.1	6.33	2	0.7%	-0.45 [-1.98, 1.08]
Bedell, 1976	20.7	5.98	10	26.1	6.33	3	0.9%	-0.83 [-2.18, 0.51]
Chang-Liang 1976	68.45	27.58	20	86.78	31.94	6	1.2%	-0.62 [-1.55, 0.31]
Chang-Liang 1976	79.91	28.98	21	86.78	31.94	6	1.2%	-0.23 [-1.13, 0.68]
Chang-Liang 1976	66.63	26.64	23	86.78	31.94	5	1.1%	-0.71 [-1.70, 0.28]
Counts 1978	24.2	6.88	10	24.15	3.74	10	1.2%	0.01 [-0.87, 0.89]
Counts 1978	18.8	7.98	10	24.15	3.74	10	1.2%	-0.82 [-1.74, 0.10]
Dawley 1973	209.17	29.03	12	238	33.78	24	1.4%	-0.87 [-1.60, -0.15]
Deffenbacher 1979a	22.77	7.41	15	36.08	6.87	17	1.3%	-1.82 [-2.66, -0.98]
Deffenbacher 1979a	23.16	6.68	17	36.08	6.87	18	1.3%	-1.86 [-2.67, -1.05]
Deffenbacher 1979b	16.41	6.7	9	28.2	5.2	5	0.9%	-1.77 [-3.11, -0.43]
Deffenbacher 1979b	20.36	7.92	9	28.2	5.2	5	1.0%	-1.03 [-2.21, 0.15]
Deffenbacher 1980a	24.92	7.89	14	30.72	5.51	15	1.3%	-0.83 [-1.60, -0.07]
Deffenbacher 1980b	28.4	5.49	15	36.85	5.63	16	1.3%	-1.48 [-2.29, -0.67]
Deffenbacher 1980b	29.46	8.03	16	36.85	5.63	15	1.3%	-1.03 [-1.79, -0.28]
Denney 1977	53.8	29.8	13	91.78	28.41	10	1.2%	-1.25 [-2.17, -0.34]
Denney 1977	53.8	38.8	9	91.78	28.41	10	1.1%	-1.08 [-2.06, -0.10]
Denney 1977	57.3	25	14	91.78	28.41	10	1.2%	-1.26 [-2.16, -0.36]
Denney 1977	51.8	35	11	91.78	28.41	9	1.1%	-1.19 [-2.16, -0.22]
Finger 1977	30.02	6.93	12	35.49	5.4	4	1.0%	-0.78 [-1.95, 0.39]
Gallagher 1978	20.2	7.18	12	25.2	5.28	24	1.4%	-0.82 [-1.54, -0.10]
Guidry 1975	128.08	25.67	12	165.5	32.02	24	1.3%	-1.22 [-1.97, -0.46]
Holahan 1979	18.67	7.76	15	23.08	6.05	6	1.1%	-0.58 [-1.54, 0.39]
Horne 1977	18.9	5.07	20	21.35	4.28	5	1.1%	-0.48 [-1.47, 0.51]
Horne 1977	11.05	5.07	20	21.35	4.28	5	1.0%	-2.02 [-3.17, -0.86]
Horne 1977	17.05	3.01	20	21.35	4.28	5	1.1%	-1.27 [-2.33, -0.22]
Kostka 1974	35	17.89	6	28.85	20.42	6	1.0%	0.30 [-0.85, 1.44]
Kostka 1974	51.6	12.18	5	28.85	20.42	6	0.9%	1.21 [-0.14, 2.55]
Levine 1980	5.07	1.12	15	5.12	1.36	9	1.3%	-0.04 [-0.87, 0.79]
Levine 1980	4.15	1.41	13	5.12	1.36	10	1.2%	-0.67 [-1.52, 0.18]
Levine 1980	4.78	0.8	14	5.12	1.36	9	1.3%	-0.31 [-1.16, 0.53]
Lomont 1971	32.5	8.17	8	35.47	6.61	17	1.3%	-0.40 [-1.25, 0.45]
Lurie 1980	17.33	6.25	6	25.57	7.25	5	0.9%	-1.12 [-2.45, 0.20]
Lurie 1980	21.5	4.89	8	25.57	7.25	4	0.9%	-0.66 [-1.90, 0.58]
Maxfield 2000	46.71	15.16	7	56	10.14	8	1.1%	-0.69 [-1.74, 0.37]
Meichenbaum 1972	26	4.1	8	30	4.6	3	0.8%	-0.87 [-2.27, 0.53]
Melnick 1976	134.89	28.52	9	163.67	19.97	18	1.2%	-1.21 [-2.08, -0.34]
Mitchell 1970	28.45	3	31	32.42	3.87	6	1.2%	-1.24 [-2.16, -0.31]
Prochaska 1971	-2	7.72	13	-1.66	9.73	7	1.2%	-0.04 [-0.96, 0.88]
Prochaska 1971	-0.23	4.19	13	-1.66	9.73	8	1.2%	0.20 [-0.68, 1.09]
Prochaska 1971	-3.67	10.23	13	-1.66	9.73	8	1.2%	-0.19 [-1.08, 0.69]
Raijah 2014	20.2	5	20	39.5	4	20	1.0%	-4.18 [-5.32, -3.03]
Reed 1980	28.8	7	9	31.85	5.98	18	1.3%	-0.47 [-1.28, 0.34]
Ricketts 1984	86.54	26.84	13	67.85	20.76	4	1.0%	0.69 [-0.46, 1.84]
Romano 1978	109.2	30.11	10	167.8	27.6	3	0.7%	-1.84 [-3.38, -0.29]
Romano 1978	127.6	38.4	10	167.8	27.6	4	0.9%	-1.05 [-2.29, 0.20]
Romano 1978	116.2	23.33	10	167.8	27.6	3	0.7%	-1.99 [-3.57, -0.41]
Russell 1976	7	4.9	5	13.7	3.6	3	0.7%	-1.29 [-2.98, 0.39]
Russell 1976	9.7	4.2	7	13.7	3.6	4	0.9%	-0.91 [-2.23, 0.41]
Russell 1982	8.2	3.9	10	11.4	4.05	6	1.1%	-0.77 [-1.82, 0.29]
Russell 1982	9.8	4	9	11.4	4.05	7	1.1%	-0.38 [-1.38, 0.62]
Russell 1982	7.8	5.8	6	11.4	4.05	7	1.0%	-0.68 [-1.81, 0.45]
Sapp 1996	52.3	8.63	15	48.58	6.4	15	1.4%	0.48 [-0.25, 1.20]
Saravanan 2014	19.28	5.56	32	38.54	4.29	33	1.3%	-3.84 [-4.68, -3.00]
Subtotal (95% CI)			**771**			**565**	**64.9%**	**-0.83 [-1.06, -0.60]**
Heterogeneity: Tau² = 0.55; Chi² = 192.17, df = 58 (P < 0.00001); I² = 70%								
Test for overall effect: Z = 7.00 (P < 0.00001)								
1.1.2 CBT								
Barabasz 1981	3.71	3.2	18	7.81	3.2	9	1.2%	-1.24 [-2.12, -0.37]
Crowley 1986	45.91	14.02	31	55.5	12.94	13	1.4%	-0.69 [-1.35, -0.02]
Crowley 1986	47.33	9.84	29	55.5	12.94	14	1.4%	-0.73 [-1.39, -0.08]
D'Alelio 1981	48	12.5	18	56.6	11.5	10	1.3%	-0.69 [-1.48, 0.11]
D'Alelio 1981	42.5	9.9	17	56.6	11.5	10	1.2%	-1.30 [-2.17, -0.43]
Decker 1981	21	8.7	9	24.3	8.2	4	1.0%	-0.36 [-1.55, 0.83]
Finger 1977	28.84	8.85	11	35.49	5.4	4	1.0%	-0.76 [-1.95, 0.42]
Finger 1977	29.98	3.78	13	35.49	5.4	4	0.9%	-1.26 [-2.48, -0.04]
Holahan 1979	18.33	5.48	12	23.08	6.05	5	1.0%	-0.80 [-1.89, 0.29]
Meichenbaum 1972	24.5	3.8	8	30	4.6	2	0.7%	-1.27 [-2.98, 0.44]
Orbach 2007	47.31	9.49	30	54.25	11.31	28	1.5%	-0.66 [-1.19, -0.13]
Reiss 2017	56.8	13.08	42	57.03	13.36	24	1.6%	-0.02 [-0.52, 0.48]
Reiss 2017	56.3	10.52	48	57.03	13.36	24	1.6%	-0.06 [-0.55, 0.43]
Ricketts 1984	73.4	30.81	15	67.85	20.76	5	1.1%	0.18 [-0.83, 1.20]
Subtotal (95% CI)			**301**			**156**	**16.9%**	**-0.58 [-0.83, -0.33]**
Heterogeneity: Tau² = 0.06; Chi² = 18.50, df = 13 (P = 0.14); I² = 30%								
Test for overall effect: Z = 4.47 (P < 0.00001)								
1.1.3 SST								
Altmaier 1981	52.54	12.85	11	54	7.73	4	1.0%	-0.12 [-1.26, 1.03]
Barabasz 1981	5.93	3	18	7.81	3.2	9	1.3%	-0.59 [-1.41, 0.22]
Decker 1981	18.1	9.5	8	24.3	8.2	4	0.9%	-0.63 [-1.87, 0.61]
Hahm 2016	51.9	9.88	43	48.43	9.9	46	1.6%	0.35 [-0.07, 0.77]
Horne 1977	22.1	3.27	20	21.35	4.28	5	1.1%	0.21 [-0.77, 1.19]
Lent 1978	24.9	5.7	10	26.3	5.6	9	1.2%	-0.24 [-1.14, 0.67]
Lurie 1980	25.37	5.57	8	25.57	7.25	5	1.0%	-0.03 [-1.15, 1.09]
McCordick 1981	-1.11	4.57	9	0.22	4.35	3	0.9%	-0.27 [-1.58, 1.04]
Mitchell 1975	32.8	2.4	20	32.2	1.6	4	1.1%	0.25 [-0.83, 1.33]
Ricketts 1984	60.6	17.91	20	67.85	20.76	4	1.1%	-0.38 [-1.46, 0.70]
Subtotal (95% CI)			**167**			**93**	**11.2%**	**0.02 [-0.25, 0.28]**
Heterogeneity: Tau² = 0.00; Chi² = 6.98, df = 9 (P = 0.64); I² = 0%								
Test for overall effect: Z = 0.12 (P = 0.90)								
1.1.4 Combined								
Altmaier 1981	44.36	8.15	11	54	7.73	4	0.9%	-1.13 [-2.36, 0.11]
Lent 1978	15.7	6.9	11	26.3	5.6	8	1.1%	-1.58 [-2.65, -0.51]
Lent 1978	18.4	6.6	10	26.3	5.6	9	1.1%	-1.23 [-2.23, -0.23]
McCordick 1981	-1.9	4.58	10	0.22	4.35	2	0.7%	-0.43 [-1.96, 1.10]
McCordick 1981	-5.38	4.14	8	0.22	4.35	2	0.7%	-1.21 [-2.91, 0.48]
McCordick 1981	-6.78	5.45	9	0.22	4.35	2	0.7%	-1.20 [-2.85, 0.46]
Mitchell 1975	30.9	1.5	24	32.2	1.6	3	0.9%	-0.84 [-2.06, 0.39]
Mitchell 1975	26	1.8	40	32.2	1.6	3	0.8%	-3.40 [-4.79, -2.00]
Subtotal (95% CI)			**123**			**33**	**7.0%**	**-1.38 [-1.96, -0.81]**
Heterogeneity: Tau² = 0.24; Chi² = 10.76, df = 7 (P = 0.15); I² = 35%								
Test for overall effect: Z = 4.72 (P < 0.00001)								
Total (95% CI)			**1362**			**847**	**100.0%**	**-0.76 [-0.93, -0.58]**
Heterogeneity: Tau² = 0.43; Chi² = 267.65, df = 90 (P < 0.00001); I² = 66%								
Test for overall effect: Z = 8.56 (P < 0.00001)								
Test for subgroup differences: Chi² = 30.93, df = 3 (P < 0.00001), I² = 90.3%								



2.3.4.3 Efficacy of interventions for reducing test anxiety, at follow-up

Fifteen studies reported follow-up data collection. However, data were only available for 12 studies (studies 1, 4, 6-7, 21-22, 24-25, 32, 38-39, 44), with three studies not reporting sufficient data to be included in the analyses (studies 18-19, 29). Only data from the first follow-up period (M = 10.0 weeks, SD = 6.8) was included in analyses (studies 21, 24-25 had a second follow-up period). Data from 294 participants across 19 treatment arms/conditions and 239 participants in control conditions were included. Of the active intervention conditions, data were available for BT (14 conditions, k = 9), CBT (4 conditions, k = 3), and SST (1 condition, k = 1). Overall, interventions for test anxiety were superior to control conditions at follow-up (g = -0.97, 95% CI -1.31 to -0.64, p < .001) (see Figure 2.4 for forest plot; see also Appendix D, Figure D.3 for annotated forest plot and Figure D.4 for the funnel plot). Significant heterogeneity was indicated (I^2 = 64%). Subgroup analysis showed that BT (g = -1.12, 95% CI -1.47 to -0.77, p < .001, k = 9) and CBT (g = -0.40, 95% CI -0.92 to 0.12, p = .007, k = 3) were superior to control conditions. Estimates for the other subgroups (SST, Combined) could not be computed due to insufficient data. Inspection of the forest plot revealed no outliers.

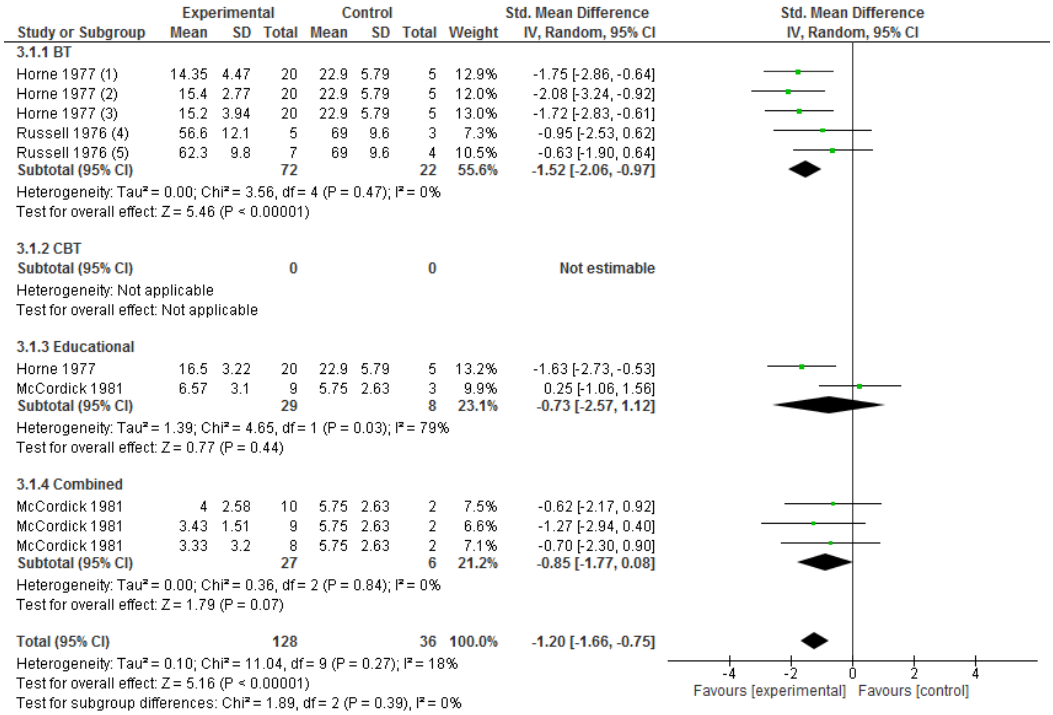
Figure 2.4. Forest plot of efficacy of interventions versus control conditions for reducing test anxiety severity (at follow-up).



2.3.4.4 Efficacy of interventions for reducing in-situ state test anxiety

Data from 10 conditions (k = 3) were analysed (from studies 13, 16, 32). In total, data from 128 participants across 10 treatment arms/conditions and 36 participants in control conditions were included in the analyses. Overall, interventions were superior to control conditions in reducing state test anxiety experienced immediately prior to 'real world' examinations (g = -1.20, 95% CI -1.66 to -0.75, p < .001, k = 3) (see Figure D.5 in Appendix D for annotated firest plot). Significant heterogeneity was not indicated (I^2 = 18%). No outliers were identified from the forest plot (see Appendix Figure D.6 for the funnel plot).

Figure 2.5. Forest plot of efficacy of interventions versus control conditions for reducing in-situ (state) test anxiety severity.

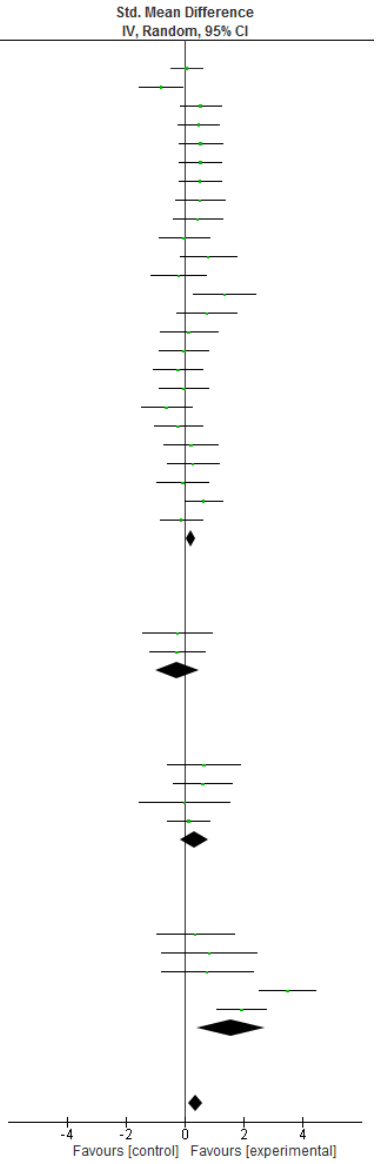


2.3.4.5 Efficacy of interventions for improving academic performance, at post-treatment

Data from 36 treatment arms (k = 17) were analysed (from studies 2-3, 8-9, 12, 14, 16, 21, 23-26, 32-33, 37-38, 41), which comprised 25 treatment arms/conditions within the BT category (k = 14), five within Combined (k = 2), four within SST (k = 4), and two conditions for CBT (k = 2). In total, data from 548 participants across active treatment arms/conditions and 363 participants in control conditions were included in the analyses (see Figure 2.6, and Appendix D Figure D.7 for the annotated forest plot, and Figure D.8 for the funnel plot).

Figure 2.6. Forest plot of efficacy of interventions versus control for improving academic performance (at post-treatment).

Study or Subgroup	Experimental Mean	Experimental SD	Experimental Total	Control Mean	Control SD	Control Total	Weight	Std. Mean Difference IV, Random, 95% CI
4.1.1 BT								
Anton 1976	2.49	0.77	32	2.43	0.75	22	3.8%	0.08 [-0.47, 0.62]
Bauman 1994	25.7	5.96	15	30.32	5.32	15	3.2%	-0.80 [-1.54, -0.05]
Deffenbacher 1979a (1)	2.76	0.56	17	2.41	0.66	17	3.4%	0.56 [-0.13, 1.25]
Deffenbacher 1979a (2)	2.73	0.59	15	2.41	0.66	18	3.3%	0.50 [-0.20, 1.19]
Deffenbacher 1980a	3.07	0.73	14	2.6	0.91	15	3.2%	0.55 [-0.19, 1.30]
Deffenbacher 1980b	2.5	0.73	16	2.03	0.92	15	3.3%	0.55 [-0.17, 1.27]
Deffenbacher 1980b	2.47	0.64	15	2.03	0.92	16	3.3%	0.54 [-0.18, 1.26]
Denney 1977 (3)	2.97	0.78	14	2.56	0.66	10	3.0%	0.54 [-0.29, 1.37]
Denney 1977 (4)	2.89	0.74	13	2.56	0.66	10	3.0%	0.45 [-0.39, 1.29]
Denney 1977 (5)	2.55	0.74	11	2.56	0.66	10	2.9%	-0.01 [-0.87, 0.84]
Denney 1977 (6)	3.07	0.53	9	2.56	0.66	9	2.6%	0.81 [-0.16, 1.78]
Holahan 1979	2.87	0.75	15	3.01	0.58	6	2.7%	-0.19 [-1.14, 0.76]
Horne 1977 (7)	3.09	0.77	20	2.05	0.58	5	2.4%	1.36 [0.30, 2.42]
Horne 1977 (8)	2.65	0.79	20	2.05	0.58	5	2.5%	0.77 [-0.24, 1.77]
Horne 1977 (9)	2.15	0.65	20	2.05	0.58	5	2.6%	0.15 [-0.83, 1.13]
Levine 1980 (10)	2.89	0.62	13	2.9	0.52	10	3.0%	-0.02 [-0.84, 0.81]
Levine 1980 (11)	2.76	0.72	14	2.9	0.52	9	3.0%	-0.21 [-1.05, 0.63]
Levine 1980 (12)	2.89	0.62	15	2.9	0.52	9	3.0%	-0.02 [-0.84, 0.81]
Lomont 1971	59.13	6.69	8	65.35	10.92	17	2.9%	-0.61 [-1.47, 0.25]
Melnick 1976	69.89	12.55	9	72.56	12.94	18	3.1%	-0.20 [-1.00, 0.60]
Prochaska 1971 (13)	0.35	2.89	13	-0.29	1.87	7	2.7%	0.24 [-0.69, 1.16]
Prochaska 1971 (14)	0.26	1.7	13	-0.29	1.87	8	2.8%	0.30 [-0.59, 1.19]
Prochaska 1971 (15)	-0.45	3.89	13	-0.29	1.87	8	2.8%	-0.05 [-0.93, 0.83]
Raijah 2014	2.9	0.3	20	2.7	0.3	20	3.5%	0.65 [0.02, 1.29]
Sapp 1996	54.19	2.58	15	54.63	4.73	15	3.3%	-0.11 [-0.83, 0.60]
Subtotal (95% CI)			**379**			**299**	**75.3%**	**0.22 [0.05, 0.40]**
Heterogeneity: Tau² = 0.04; Chi² = 29.34, df = 24 (P = 0.21); I² = 18%								
Test for overall effect: Z = 2.46 (P = 0.01)								
4.1.2 CBT								
Decker 1981	1.58	0.78	9	1.79	0.89	4	2.1%	-0.24 [-1.42, 0.94]
Holahan 1979	2.88	0.47	12	3.01	0.58	7	2.7%	-0.24 [-1.18, 0.69]
Subtotal (95% CI)			**21**			**11**	**4.9%**	**-0.24 [-0.98, 0.49]**
Heterogeneity: Tau² = 0.00; Chi² = 0.00, df = 1 (P = 1.00); I² = 0%								
Test for overall effect: Z = 0.65 (P = 0.52)								
4.1.3 Educational								
Decker 1981	2.19	0.31	8	1.79	0.89	4	2.0%	0.67 [-0.58, 1.91]
Horne 1977	2.55	0.8	20	2.05	0.58	5	2.6%	0.63 [-0.37, 1.63]
McCordick 1981	-0.12	1.31	9	-0.13	0.58	2	1.6%	0.01 [-1.52, 1.54]
Mitchell 1975	19.8	18.9	20	16.63	21.79	12	3.3%	0.15 [-0.56, 0.87]
Subtotal (95% CI)			**57**			**23**	**9.4%**	**0.34 [-0.16, 0.84]**
Heterogeneity: Tau² = 0.00; Chi² = 1.03, df = 3 (P = 0.79); I² = 0%								
Test for overall effect: Z = 1.34 (P = 0.18)								
4.1.4 Combined								
McCordick 1981	0.2	0.84	9	-0.13	0.58	3	1.9%	0.38 [-0.94, 1.70]
McCordick 1981	0.71	0.94	8	-0.13	0.58	2	1.5%	0.84 [-0.78, 2.46]
McCordick 1981	0.5	0.76	10	-0.13	0.58	2	1.5%	0.78 [-0.78, 2.35]
Mitchell 1975	56.5	6	40	16.63	21.79	11	2.6%	3.50 [2.53, 4.48]
Mitchell 1975	43.5	6.6	24	16.63	21.79	12	3.0%	1.94 [1.10, 2.78]
Subtotal (95% CI)			**91**			**30**	**10.4%**	**1.58 [0.41, 2.76]**
Heterogeneity: Tau² = 1.38; Chi² = 19.12, df = 4 (P = 0.0007); I² = 79%								
Test for overall effect: Z = 2.64 (P = 0.008)								
Total (95% CI)			**548**			**363**	**100.0%**	**0.37 [0.14, 0.61]**
Heterogeneity: Tau² = 0.30; Chi² = 90.17, df = 35 (P < 0.00001); I² = 61%								
Test for overall effect: Z = 3.12 (P = 0.002)								
Test for subgroup differences: Chi² = 6.86, df = 3 (P = 0.08), I² = 56.3%								



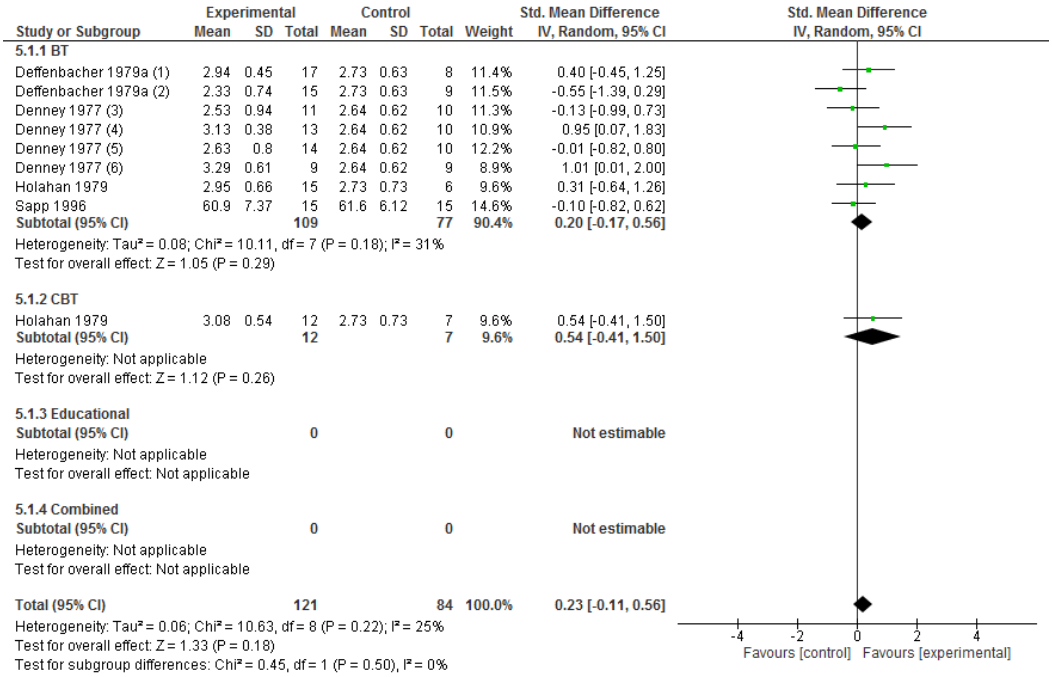
Results from single examinations came from 10 treatment arms/conditions (k = 4), while GPA results came from 26 treatment arms/conditions (k = 13). Overall, treatments were superior to control conditions in improving academic performance producing a small-to-moderate effect (g = 0.37, 95% CI 0.14 to 0.61, p = .002, k = 36).

Significant heterogeneity was indicated (I^2 = 61%). Subgroups analysis showed that BT (g = 0.22, 95% CI 0.05 to 0.40, p = .021, k = 13) and Combined interventions (g = 1.58, 95% CI 0.41 to 2.76, p < 001, k = 2) were superior to control conditions, but SST (g = 0.34, 95% CI -0.16 to 0.84, p = .790, k = 4) and CBT (g = -0.24, 95% CI -0.98 to 0.49, p = .520, k = 4) were not significantly different from control conditions. Inspection of the forest plot revealed one outlier in the Combined subgroup (systematic desensitization and SST intervention of Mitchell et al., 1975). Removal of this outlier resulted the overall effect estimate changing to g = 0.28 (95% CIs 0.11 to 0.46) and the Combined interventions subgroup effect size changing to g = 1.15 (95% CIs 0.33 to 1.96).

2.3.4.6 Efficacy of interventions for improving academic performance, at follow-up

Four studies reported follow-up data collection (studies 14, 21, 23, 38). Only data from the first follow-up period (M = 6.8 weeks, SD = 2.2) was included (one study had a second follow-up period). It was not possible to calculate the mean follow-up period as insufficient data was provided. In total data from 121 participants across nine treatment arms/conditions and 84 participants in control conditions were included in analyses. Of the active intervention conditions, data were available for BT (8 conditions, k = 4) and CBT (1 condition, k = 1). Overall, interventions for test anxiety were not significantly different to control conditions at follow-up (g = 0.23, 95% CI -0.11 to 0.56, p = .180) (see Appendix D, Figure D.9 for annotated forest and Figure D.10 for the funnel plot). Significant heterogeneity was not indicated (I^2 = 25%). No outliers were identified.

Figure 2.7. Forest plot of efficacy of interventions versus control conditions for improving academic performance (at follow-up).



A summary of effect sizes across intervention approaches is in Table 2.3.

2.3.4.7 Moderators of intervention effect

Only moderators of treatment outcome for BT interventions at post-treatment for test anxiety reduction were examined, as there was insufficient data to examine moderators for the other treatment approaches (and insufficient data for all approaches at follow-up). Significant heterogeneity was indicated within outcomes for specific BT interventions (I^2 = 70%). The two sample-dependent moderators – gender, pre-treatment test anxiety severity – had to be dropped the analyses due to insufficient data. Thus, only format of the intervention (group vs. individual), dosage (number of treatment hours), and manualization (poor vs. adequate) were entered into the meta-regression; none were significant

moderators of treatment outcome. Results from the multiple meta-regression analyses are presented in Table 2.4.

Table 2.3. Summary of meta-analytic findings (active interventions vs. control conditions) based upon Hedges' *g*.

	Test anxiety			**Academic performance**	
Intervention	**Post**	**Follow-up**	**State**	**Post**	**Follow-up**
BT	-0.83 (-1.06, -0.60)	-1.15 (-1.38, -0.91)	-1.52 (-2.06, -0.97)	0.22 (0.05, 0.40)	0.20 (-0.17, -0.56)
CBT	-0.58 (-0.83, -0.33)	-0.31 (-0.64, 0.02)	-	-0.24 (-0.98, 0.49)	0.54 (-0.41, 1.50)
SST	0.02 (-0.25, 0.28)	-0.95 (-2.24, 0.34)	-0.73 (-2.57, 1.12)	0.34 (-0.16, 0.84)	-
Combined	-1.38 (-1.96, -0.81)	-	-0.85 (-1.77, 0.08)	1.58 (0.41, 2.76)	-
Overall	*-0.76* *(-0.93, -0.58)*	*-0.87* *(-1.06, -0.68)*	*-1.20* *(-1.66, -0.75)*	*0.37* *(0.14, 0.61)*	*0.23* *(-0.11, 0.56)*

Notes.

Test anxiety: negative effect sizes indicate reduction in test anxiety; positive effect sizes indicate increase in test anxiety. Academic performance: negative effect sizes indicate decrease in academic performance; positive effect sizes indicate increase in academic performance. BT = Behavioural therapy, CBT = Cognitive-behavioural therapy, Combined = Combined psychotherapy (BT, CBT) and SST; SST= Study Skills Training.

Table 2.4. Moderators of outcome at post-treatment for behavioural therapies.

Regression coefficients	***b***	***SE***	***LL***	***UL***	***Z***	***p***
Constant	-0.28	0.37	-1.01	0.45	-0.75	.451
Delivery (group vs. individual)	-0.20	0.32	-0.82	0.42	-0.63	.527
Dosage (no. of hours)	-0.09	0.05	-0.18	0.00	-1.87	.062
Country (USA vs. other)	-0.64	0.41	-1.44	0.15	-1.58	.114
Manualization (poor vs. adequate)	0.31	0.25	0.21	0.80	1.27	.205

Notes.

R^2 = 5.58%. *LL* = (95% confidence interval) Lower Limit; UL = (95% confidence interval) Upper Limit. USA = United States of America.

2.3.4.8 Sensitivity analyses

Sensitivity analyses were planned to evaluate the robustness of the results by removing studies with a high risk of bias and comparing the results when all studies are included. However, these analyses could not be conducted as the poor quality of reporting was such that the overwhelming majority of risk of bias domains were classified as unclear (see Figure 2.2) and so no RCT could be considered of high (or low) risk of bias.

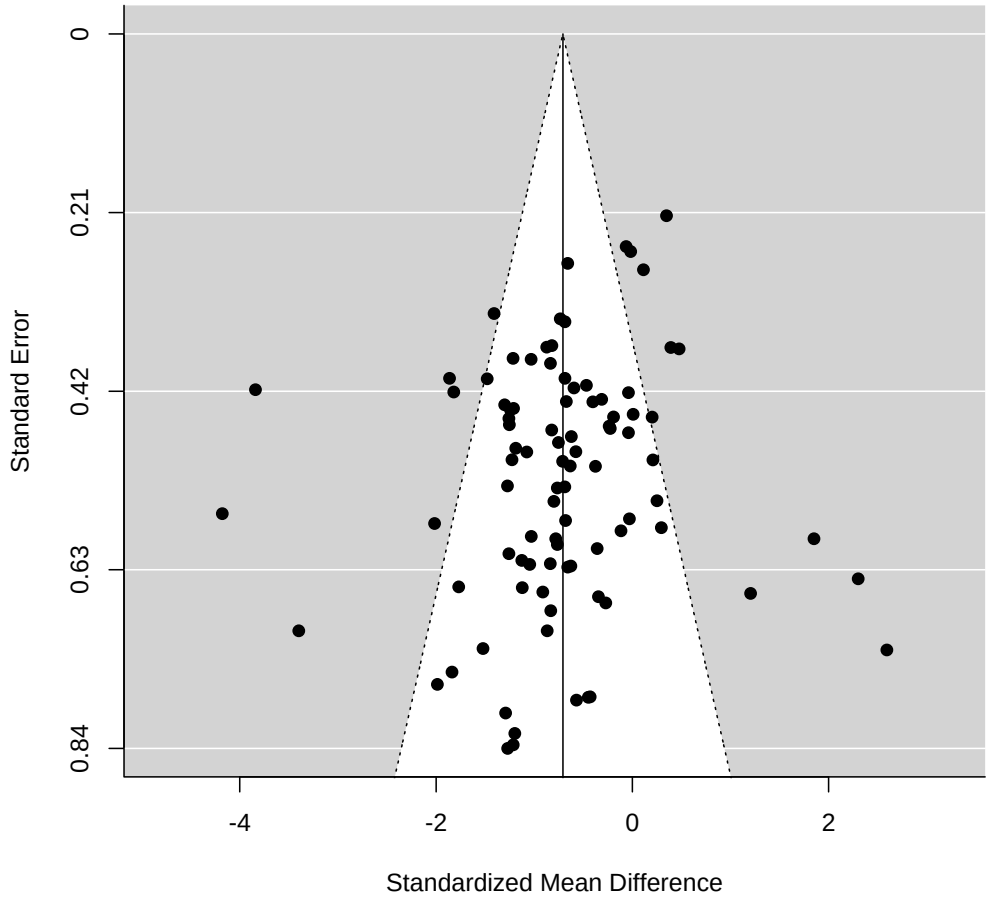
2.3.4.9 Assessment of publication bias

Publication bias was explored using data derived from test anxiety interventions at post-treatment. Inspection of the funnel plot (see Figure 2.8) revealed minor asymmetry and this was confirmed by Egger's regression test (z = -2.11, p = .003). More noticeable from the funnel plot was the absence of studies with very precise estimates (i.e., that had sufficiently large samples) included within the analysis.

2.3.5 Quality of evidence

GRADE assessments were made for each intervention subgroup across five outcomes: reducing test anxiety at (i) post-treatment, (ii) follow-up; (iii) reducing in-situ or state test anxiety experienced immediately prior to sitting a real examination; and improving academic performance at (iv) post-treatment, and (v) follow-up (see Appendix E, online only). For a GRADE assessment to be made for an intervention approach on any of the above outcomes there must be evidence from at least two independent studies.

Figure 2.8. Funnel plot of intervention effects and study precision (standard errors) for reducing test anxiety severity (at post-treatment).



The overall quality of evidence was poor, with 'low quality' being the highest grading awarded for any intervention on any outcome. The main reasons for the low-quality ratings were: (i) potential for risk of bias, indicated by nearly 80% of ratings being scored 'unclear', (ii) inconsistency as indicated by I^2 scores, and (iii) the relative imprecision of the effect size estimates (i.e., relatively large confidence intervals around the effect size estimate). There were also large gaps in the evidence base for all intervention approaches except for BTs. For reducing test anxiety at post-treatment, the evidence was rated as 'low

quality' for all intervention approaches. At post-treatment BT and CBT were both rated as 'low quality', while lack of evidence meant no ratings could be given to SST and Combined intervention approaches. Evidence for reducing in-situ or state test anxiety was rated as 'very low quality' for BT and SST, but no ratings could be made for CBT and Combined approaches. With regard to improving academic performance at post-treatment, evidence for BT, CBT, SST was rated as 'low quality', while evidence for Combined interventions approaches were rated as 'very low quality'. For improving academic performance at follow-up, the evidence for BT was rated as 'very low quality'; there was insufficient evidence for the other approaches to award a rating. Overall, these findings indicate that future studies will very likely have significant impact on the effect size estimates and their confidence intervals.

2.4. Discussion

2.4.1 Summary of main results

This meta-analysis is the first to systematically examine evidence regarding the efficacy of interventions for test-anxious university undergraduate students. The aims of this meta-analysis were to examine the relative efficacy of interventions for reducing test anxiety and improving academic performance. The final aim was to also examine possible moderators of intervention effect. Overall, interventions were superior to control conditions at post-treatment for reducing test anxiety, with the summary effect size across all interventions (g = -0.76) indicating a moderate-to-large effect (based upon Cohen's 1988 guidelines). However, as expected, there was significant heterogeneity between study outcomes. In a subgroup analysis, combined psychological and SST

interventions, BT, and CBT were superior to control conditions, but SST delivered alone was not superior to control conditions. An important question is whether intervention effects were sustained over follow-up (compared with control conditions)? Overall, interventions were superior to control conditions at follow-up, with a large effect size found (g = -0.97). However, the length of the follow-up period across the 12 studies that reported data was relatively brief (at a mean of 9.4 weeks) and did not specify when follow-up assessment took place in relation to the timing of examinations. Only BT was superior to control conditions at follow-up.

Given test anxiety is a situation-specific form of anxiety, interventions had a beneficial impact on anxiety experienced immediately before an examination was investigated (in-situ state anxiety). A large overall effect size was found (g = -1.20) indicating interventions may help alleviate test anxiety immediately prior to an examination. However, only the effect size estimate was based on data from just three studies.

Interventions also improve academic performance. Interventions were significantly superior control conditions (g = 0.37) at post-treatment. Subgroup analysis found BT and Combined psychological and educational intervention approaches superior to control conditions but SST delivered alone was not superior to control conditions. There was insufficient data available to calculate an effect for CBT. At follow-up, no significant differences between active interventions and control conditions were found, though data were derived from just four studies.

The moderation analysis of BT effect sizes at post-treatment, which examined delivery format (group vs. individual), dosage (number of contact

hours), and whether the intervention was manualized, found all were non-significant predictors. There was insufficient data for moderator analyses to be conducted for the other interventions and follow-up periods.

Although there were generally large effect sizes for treatments compared to control conditions, the overall quality of evidence was poor – or insufficient details were provided by the individual RCTs meaning GRADE assessments could not be made. Across the intervention approaches, BT had the most substantive evidential support. However, even within the BT approach there was high heterogeneity that suggested different levels of efficacy between specific behavioural treatments. There is less evidence for CBT with moderate post-treatment effects but limited follow-up data negates assessment of longer-term effects. SST interventions do not appear effective for reducing test anxiety nor are particularly effective at improving academic performance. However, SSTs interventions included were generally poorly described, and it might be that manualized educational interventions, informed by recent literature in this field (e.g., retrieval-practice effect; Karpicke & Roediger, 2008) would deliver superior outcomes. It should also be noted that the primary aim of the SST interventions here was to decrease test anxiety with improvement in academic performance a secondary aim, whereas most educational interventions aim to improve academic performance first with the impact on test anxiety rarely considered. Combined psychological and SST interventions seem promising. However, a lack of data meant an effect size could only be calculated at post-treatment, so the longer-term efficacy of these approaches is unknown.

2.4.2 Overall completeness and applicability of evidence

Studies included in this review are a partial representation of the studies conducted and published. Seventy-eight eligible studies were identified. Thirty-four of these studies could not be included within the meta-analyses as insufficient data were reported in the study manuscripts and, when contacted, authors no longer had access to the data. Additionally, publication bias was indicated by marginal funnel plot asymmetry, suggesting some trials with poorer outcomes have not been published in peer-reviewed journals. Science is a collaborative and cumulative endeavour and efforts should be made in the future to catalogue and preserve such trial data in the future (Goldacre, 2015).

All reviewed studies included university undergraduates only, but important demographic and clinical information was frequently not reported. Only 27 of the 44 studies reported the percentage of male and female participants in their samples, with the ratio of male-to-female participants varying widely, from 0% (Prochaska, 1971) to 97% (Bauman, 1994). Just one study reported the proportion of male and female participants for intervention conditions (Reiss et al., 2017). These figures cast doubt on the representativeness of the samples used in some studies.

Reporting of the clinical characteristics of study samples was similarly incomplete. Only 30 of the studies used cut-off criteria, with the other 14 studies just enrolling self-selected volunteers who responded to advertisements. The most frequent method was to screen an entire year's cohort and offer participation in the trial to those scoring above a cut-score or set percentile. However, there were large discrepancies in how the cut-offs were applied. For example, eligible participants could come from the top 50% of scores (Melnick

& Russell, 1976) to the top 15% (Denney & Rupert, 1977). Only one study (Reiss et al., 2017) used a structured clinical interview to screen for test anxiety. Thus, the severity of pre-treatment test anxiety varied widely from study-to-study.

2.4.3 Quality of evidence

Overall, results should be interpreted with caution due the issues identified. The principal problem was the poor quality of reporting, with many details either partially reported or not reported at all. For example, only two studies reported the flow of participants from recruitment to follow-up periods (Raijha, 2014; Reiss et al., 2017). Nearly 80% of risk of bias items were scored as unclear due to insufficient information reported in study manuscripts and therefore it is not possible to be confident those trials were conducted rigorously. It also meant it was not possible to identify any high (or low) quality trials. Poor quality reporting does not necessarily mean methodologically poor trials. Rather, it means one's confidence in the estimates obtained is undermined. Most interventions examined in studies were insufficiently described, with nearly half of studies (42%) not referencing a manual (or how different components were combined). Manualization enhances the internal validity of the trial and allows replicability (Temple et al., 2018).

The review identified 44 studies. No effect size was estimated with precision for any of the subgroups, with the smallest interval between upper and lower 95% bounds of the effect size estimate being 0.46 (for BT subgroup for reducing test anxiety severity at post-treatment). This is a result of the low statistical power for individual studies, confirmed by visual inspection of the

funnel plot, where no studies are located at the apex of the funnel (i.e., this is where studies with small standard errors – indicative of large samples – should be). To detect an effect size of 0.70, with 80% power, and α of .05, a minimum of 34 participants are required for each treatment arm (two-side significance). However, only 5% (8 of 152) treatment arms had 34 or more participants, with a mean of just 15.8 participants allocated to active conditions.

Another limitation of the studies included was the large number of different self-rating scales that were used to assess test anxiety severity. This makes it difficult to compare test anxiety severity. A 'gold standard' outcome measure is needed.

2.4.4 Limitations of this review

There are several limitations to this review. Firstly, the search protocol was limited to English-language peer-reviewed published studies. Cochrane review protocols suggest attempting to find all non-published and non-English documents (Cochrane, 2011), and the inclusion of such studies would have been preferable.

Secondly, statistical analyses dealt with studies with multiple treatment arms by dividing the sample size of the control group by the number of treatment arms. This meant each participant was only included once in each analysis. Other methods for dealing with this include combining treatment arms that are conceptually similar. However, as this was the first review of interventions for test anxiety in university students, and it was felt it was better to show the results from each treatment arm, rather than obfuscating what was originally examined by combining interventions. It also decided (*a priori*) to keep the subgroups

relatively broad (e.g., BT) as it was anticipated that evidence available for interventions based upon specific models (e.g., Ellis' Rational Emotive Behaviour Therapy) would be limited.

2.4.5. Conclusion

BT is the current first choice intervention for reducing test anxiety and, in turn, improving academic performance. There is more substantial evidence for the efficacy of BT immediately following treatment and over the follow-up period compared to other intervention approaches. However, there was inconsistency in effects across the various forms of BT interventions making it difficult to recommend a specific behavioural intervention. Combined interventions (BT or CBT plus SST) are promising but there is lack of evidence for longer-term efficacy. CBT appears to be efficacious for reducing test anxiety severity, but the magnitude of its effects is less than either BT or the Combined approaches. SST is ineffective in alleviating test anxiety in the short-term and a lack of data precludes conclusions about its longer-term efficacy, while it does not appear to significantly improve short-term academic performance**.** In general, the interventions evaluated have a greater impact on test anxiety than academic performance. Moreover, overall confidence in the results should be tempered, given the relatively small number of included studies, evidence of publication bias, relatively wide confidence intervals around effect size estimates, and the poor quality of reporting of individual RCTs.

In summary, findings at post-intervention, and when outliers were removed, interventions showed medium effects for reducing test anxiety severity and weak effects for improving academic performance. There was a lack of data

for follow-up periods and for the effects of interventions on state test anxiety to be able to draw firm conclusions. The overall quality of evidence was poor. Higher quality, larger trials are required. These studies should be adequately powered, include longer follow-ups, and adhere to modern reporting standards. Additionally, given that test anxiety is a situation-specific disorder, future trials should also consider when post- and follow-up assessments are made (i.e., ideally soon after examinations) so that the effects of interventions are clearer. This review also highlights a marked drop in the number of trials evaluating interventions designed to alleviate test anxiety.

In conclusion, existing interventions for test anxiety are not particularly effective for reducing test anxiety symptoms, nor for improving test and/or academic performance, especially when outliers are removed, and the quality of evidence is considered. Moreover, most interventions are based on behavioural approaches that principally target the somatic anxiety symptoms of test anxiety. It is also notable that none of the existing interventions specifically attempt to reduce worry, the primary feature of test anxiety. Interventions that are based on psychological models that can explain the initiation and persistence of worry, along with somatic anxiety symptoms, would offer a promising avenue for improving outcomes. The next chapter will consider such models and evaluate their suitability for being applied to test anxiety.

Chapter Three

Alternative approaches to understanding test anxiety: The Intolerance of Uncertainty model and the Self-Regulatory Executive Function model

3.1 Introduction

Most interventions included in the meta-analysis (chapter 2) on the efficacy of psychological treatments for test anxiety were behavioural approaches (e.g., systematic desensitization). These approaches principally target the somatic symptoms of test anxiety (i.e., the emotionality component). However, worry is the central feature of test anxiety. None of the interventions, including the cognitive-oriented approaches, specifically targeted worry and the mechanisms that initiate and maintain it. Rather these cognitive approaches seek to modify dysfunctional attitudes about the self, the world, and the future (Beck, Emery, & Greenberg, 1985) or challenge irrational beliefs (Ellis, 1957, 1962). Additionally, none of the other established theoretical models of test anxiety (Wine, 1971; Sarason, 1988; Speilberger & Vagg, 1995; Zeidner, 1998) that interventions may be based upon can explain the initiation and maintenance of worry. Psychological models that can account for worry and anxiety may improve outcomes in test anxiety reduction and examination performance improvement.

There are four principal models of worry and anxiety: cognitive avoidance (Borkovec, 1994; Borkovec, Alcaine, & Behar, 2004), emotional dysregulation (Mennin et al., 2002; Newman & Llera, 2011), intolerance of uncertainty (IU; Dugas et al., 1998; Hebert & Dugas, 2019), and the Self-Regulatory Executive Function (S-REF; Wells & Matthews, 1994, 1996) models. Each model was initially developed to account for the persistent worry that is the key feature of Generalized Anxiety Disorder (GAD).

The cognitive avoidance model is based upon behavioural principles, proposing that worry is acquired through classical conditioning and maintained

by operant conditioning (Mowrer, 1947). The content of worrisome thinking is future oriented about events or outcomes that may or may not happen. However, as fight or flight is not available to a hypothetical future, the cognitive avoidance model posits that worry is used as an avoidant coping strategy, which is negatively reinforced as the worst-case scenarios envisioned rarely or never actually occur. Another means by which worry is reinforced concerns the finding that worry is mostly verbal-linguistic in nature (Borkovec & Inz, 1990), whilst negative imagery is more arousing that worry (Borkovec & Hu, 1990). Cognitive avoidance theory proposes that anxious individuals switch to worry in response to negative imagery, dampening their arousal, and negatively reinforcing worry, which is supported empirically (Sibrava & Borkovec, 2006). Cognitive avoidance theory was highly influential and anxious individuals' perception of the usefulness of worry is a common thread among all models. A major issue with the model is individuals experience significant negative emotion due to worry (Llera & Newman, 2010), which runs contrary to the idea that worry enables emotional avoidance. Also, the model does not explain why individuals with severe worry find their emotions so aversive (Newman & Llary, 2011).

Emotional dysregulation models build upon the ideas of cognitive avoidance model, accepting worry performs an emotion avoidant function (Mennin et al., 2002; Newman & Llera, 2011). However, they address the question why those people with severe worry and anxiety find emotions so aversive. Emotion dysregulation models propose those with severe worry and GAD have deficits in emotional functioning, experiencing emotions more intensely, and difficulties in identifying, understanding emotions, and managing their emotions. Although GAD patient and analogue samples show greater

emotional intensity and poor emotional regulation compared to controls (Mennin et al., 2005; Mennin, McLaughlin, & Flanagan, 2009), no differences have been found between GAD patients and controls in emotional awareness (Novick-Kline et al., 2005; Decker et al., 2008).

Though there is empirical support for the cognitive avoidance and emotional dysregulation models applied to GAD, there is little supporting evidence for their applicability to other anxiety problems and disorders. The most prominent models that have been applied to both GAD and other anxiety problems and disorders are the IU and S-REF models (Boswell et al., 2013; Sellers et al., 2017). Moreover, both models account for the initiation and persistence of worry and perseverative thinking. Efficacious interventions based upon the IU and S-REF models have also been developed and applied to anxiety disorders (van der Heiden, Muris, & van der Molen, 2012; Normann & Morina, 2018). Therefore, these two models are the best candidates for also accounting for test anxiety. Descriptions of the IU and S-REF models and the evidence that supports them are detailed below.

3.2 Intolerance of Uncertainty Model

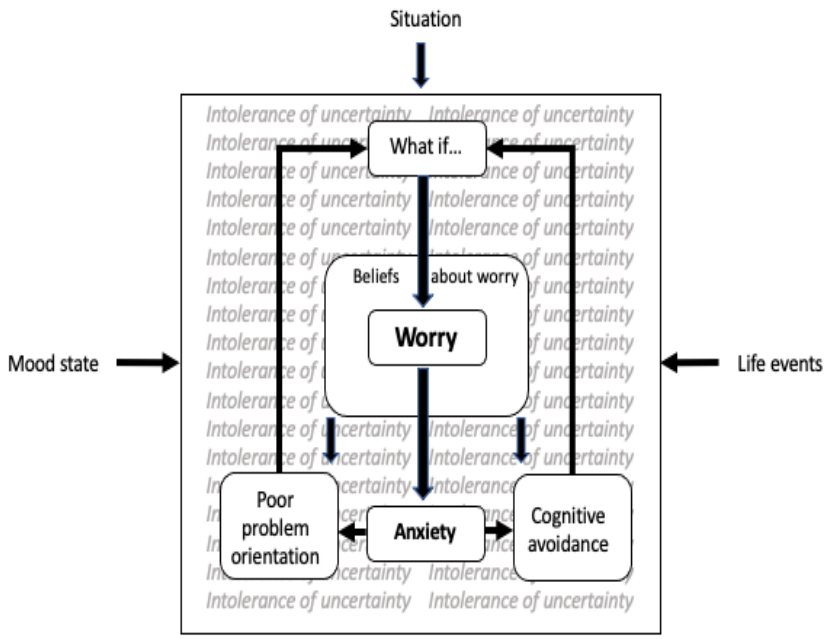
The IU model was originally developed to explain the pathological worry that was the key feature of Generalized Anxiety Disorder (GAD; Dugas et al., 1998). However, subsequently the IU model has been applied to other emotional disorders, such as social anxiety disorder (Boelen & Reijntjes, 2009) and depression (Liao & Wei, 2011). The IU model is now considered a transdiagnostic model for explaining perseverative thinking (worry, rumination) and distress in emotional problems and disorders (Boswell, Thompson-Hollands,

Farchione, & Barlow, 2013; Carleton, 2012, 2016; Mahoney & McEvoy, 2012; McEvoy & Mahoney, 2012).

The original IU model consisted of four main components: IU, cognitive avoidance, positive beliefs about worry, and negative problem orientation (see Figure 3.1; Dugas et al., 1998). This model has since been updated to also include maladaptive coping or 'safety behaviours' (Herbet & Dugas, 2019) (see Figure 3.2). IU refers to catastrophic negative beliefs about uncertainty and is central to the model (Dugas et al., 1998; Hebert & Dugas, 2019). Uncertainty is defined as an internal state of being unsure or not knowing, rather than characteristics of situations per se (Shihata et al., 2016). Individuals with high IU therefore have a marked tendency to view the chance of negative events or outcomes as intolerable and threatening, irrespective of the actual probability of the events occurring (Carleton, Norton, & Asmundson, 2007; Dugas, Gosselin & Ladouceur, 2001). Examples of IU beliefs are "uncertainty makes life intolerable", "uncertainty makes me anxious", and "uncertainty is dangerous". IU is considered non-taxonomic, with the degree of IU experienced varying along a continuum (Carlton et al., 2012).

IU acts as a 'cognitive filter' that enhances cognitive biases and leads directly to worry (Carleton et al., 2007; Dugas et al., 2001). Individuals with high IU are more likely to interpret ambiguous and uncertain situations as threatening, show facilitated engagement toward threatening stimuli over non-threatening stimuli, and attempt to remove or reduce uncertainty with worry and via seeking additional information (Dugas et al., 1998; Ladouceur, Talbot, & Dugas, 1997). IU runs in the background and therefore affects all other components of the model (Hebert & Dugas, 2019).

Figure 3.1. Original Intolerance of Uncertainty model (Dugas et al., 1998).

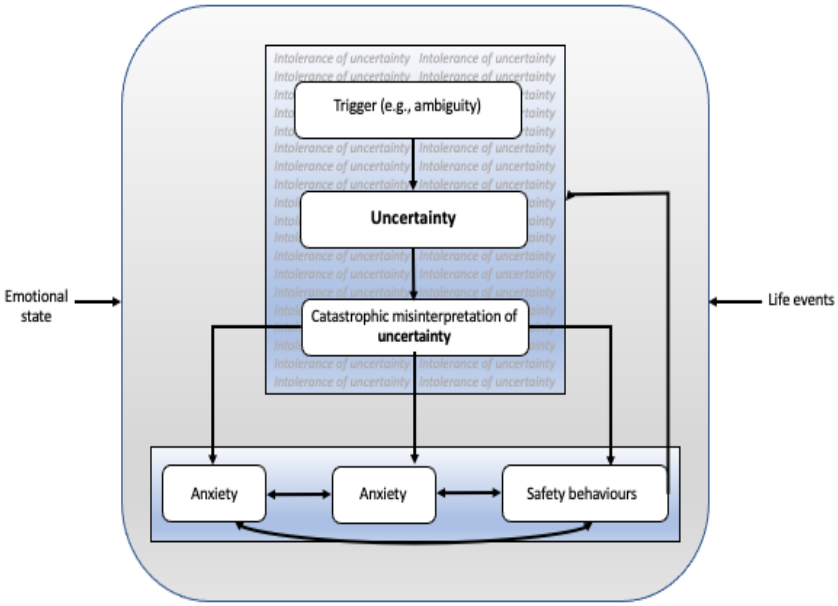


The other components of the IU model, positive beliefs about worry, negative problem orientation, and cognitive avoidance, are posited to contribute to worry via indirect pathways.

Positive beliefs about worry refer to beliefs that worry is useful for coping and problem solving, and that worry can help prevent negative events and outcomes (Dugas et al., 1998). Examples of positive beliefs about worry are "worrying helps me be prepared" and "worry helps prevent negative events from happening". Positive beliefs about worry are maintained by both positive and negative reinforcement. When a favourable outcome occurs, it is attributed to worry, i.e., positive reinforcement. However, worry is negatively reinforced when a feared event or outcome does not occur, and negative reinforcement will frequently transpire given that worst-case worry scenarios that individuals

conceive often have little chance of occurring (e.g., a student may think “I will completely bomb out on this examination” even if they are an ‘A’ grade student).

Figure 3.2. Revised Intolerance of Uncertainty model (Hebert & Dugas, 2019).



Negative problem orientation refers to dysfunctional attitudes and negative perceptions of one’s problem-solving abilities (Dugas et al., 1998). As such, individuals view problems as threats, have little confidence in their ability to solve the problems, and become easily frustrated in their problem-solving efforts. Negative problem orientation ultimately leads to limited attempts at dealing with difficult situations and can exacerbate worry.

Cognitive avoidance refers to the automatic and strategic attempts of trying to curtail or avoid distressing thoughts and mental images, such as avoiding or suppressing unwanted mental content, and is drawn from Borkovec and colleague’s cognitive avoidance theory (Borkovec, Alcaine, & Behar, 2004). Here, individuals use worry as suppressor of the physiological and emotional arousal that comes with negative thoughts and mental imagery. This has the

effect of making the suppressed cognitions implicitly take on greater salience (i.e., “I suppress it, therefore it must be dangerous”). Worry is reinforced as the negative outcomes (i.e., physiological arousal associated with the negative thoughts and imagery) are removed.

The final, and newest component, of the IU model are safety behaviours (Hebert & Dugas, 2019). Safety behaviours are maladaptive and unnecessary coping behaviours that individuals enact to reduce, prevent, or avoid a feared outcome (Clark & Wells, 1995). Safety behaviours are witnessed most acutely in individuals with obsessive-compulsive disorder (OCD), in which a thought (e.g., thinking an object is dirty) may lead to excessive behaviours (e.g., cleaning and scrubbing of the object and other things that have been in contacted with that object) to avoid the feared outcome (e.g., possibility of contamination) (Levy & Radomsky, 2016).

How would the IU model apply to test anxiety? Test anxious students view tests as threatening, primarily because concern over the uncertainty of the test, their potential performance, and the consequences of poor test performance. Positive beliefs about the usefulness of worrying, leads to overuse of worry as a means of coping. If students regard tests as problems, and doubt their ability to perform, this will also lead to worry. Finally, students may engage in test irrelevant thinking when attempting to cognitively avoid distressing imagery and thoughts.

3.2.1 Empirical support for Intolerance of Uncertainty model

There is substantive support for IU being linked to worry and emotional distress. IU is highly correlated with worry in both clinical and non-clinical

samples (Dugas et al., 1998; Dugas, Freeston, & Ladouceur, 1997; Dugas et al., 2001; Ladouceur et al., 1999). Moreover, the relationship between IU and worry is not due to shared variance with anxiety and depression (Dugas et al., 1997; Freeston et al., 1994). In experimental studies that manipulate IU, increasing the severity of IU leads to corresponding increases in the levels of worry (Ladouceur, Gosselin, & Dugas, 2000). IU is a stronger predictor of worry than positive beliefs about worry, negative problem orientation, and cognitive avoidance (Dugas et al., 2007). Changes in IU precede changes in worry and anxiety during IU-based cognitive-behavioural therapy (Dugas & Ladouceur, 2000; Dugas et al., 2003). Support for the transdiagnostic nature of IU can be found in studies showing that it is associated with a wide range of emotional issues, including GAD (Dugas et al., 1998; Counsell et al., 2017), social anxiety disorder (SAD; Boelen & Reijntes, 2009; Counsell et al., 2017), major depressive disorder (MDD; Liao & Wei, 2011), post-traumatic stress symptoms (PTSD; Oglesby et al., 2017), and obsessive-compulsive disorder (OCD; Tolin et al., 2003).

IU has been linked to biased information processing. High IU individuals are more likely to interpret ambiguous and uncertain stimuli as threatening and are more likely to recall words signifying uncertainty, compared to their low IU peers (Dugas et al., 2005). IU is associated with facilitative engagement towards uncertainty-related words and threatening visual stimuli, indicating a hypervigilance to stimuli designating uncertainty (Fergus, Bardeen, & Wu, 2013; Fergus & Carleton, 2016; Morris et al., 2018).

With regards to the other components of the IU model, there is substantive support for positive beliefs about worry being associated with worry

and anxiety (Borkovec, Hazlett-Stevens, & Diaz, 1999; Sun, Zhu, & So, 2017). Cognitive avoidance leads to the maintenance of elevated levels of worry and anxiety (Borkovec et al., 2004; Dugas et al., 1998). Negative problem orientation is associated with worry and GAD (Robichaud & Dugas, 2005a,b). Safety behaviours, such as avoidance and reassurance seeking, are associated with anxiety (Beesdo-Baum & Knappe, 2012). Currently, there is no empirical evidence linking IU with safety behaviours. Examination of components of the model show IU is most important in predicting worry severity, when partialling out the contributions of the other variables (Bottesi et al., 2016; Dugas et al., 2007).

In summary, there is considerable empirical support for central components of the IU model as applied to worry, anxiety, and emotional disorders.

3.3 The Self-Regulatory Executive Function (S-REF) model

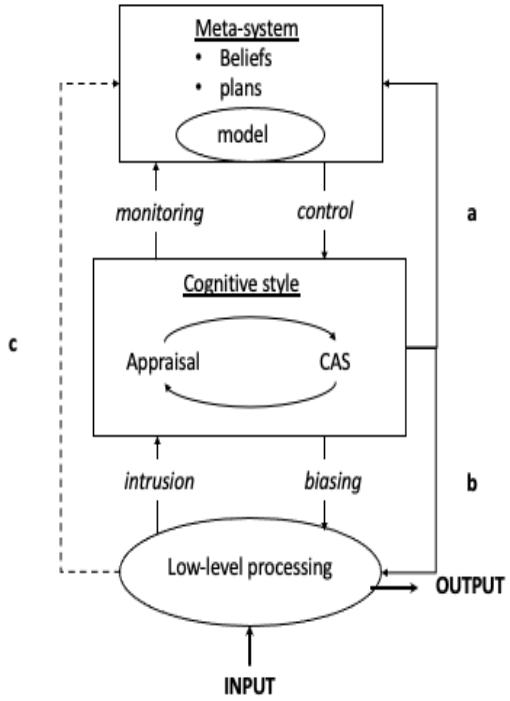
The S-REF model (Wells & Matthews, 1994, 1996) is a transdiagnostic model of psychological disorder. This model was developed due to dissatisfaction with existing psychological models and treatments, such as Beck's schema-based therapy (Beck, 1976; Beck, Emery, & Greenberg, 1985). Schema-based models failed to account for the looming maladaptive thinking style, and the factors that control this thinking style, that is present across emotional disorders. Schema-based models like Beck's assert that cognitive biases in thinking lead to psychological disorder. Cognitive biases are derived from dysfunctional attitudes about the self, the world, and the future (e.g., "I am a failure"). However, there is a conundrum in the Beckian theory in that

everyone will occasionally have negative thoughts, and sometimes even believe those thoughts, but not everyone will develop a psychological disorder or experience sustained distress. The S-REF model answers this conundrum. Here, persistent emotional disorder is not a consequence of the specific content of negative thoughts but rather it is due to recurrent and inflexible self-referent thinking styles that result from maladaptive beliefs about thinking (i.e., metacognition). It was the first model to account for cognition and metacognitive beliefs and processes involved in the top-down or voluntary control of emotional disorder (Wells, 2009).

The S-REF model specifies a cognitive architecture consisting of three interacting levels (see Figure 3.3): a low-level of automatic processing, a level of conscious or 'online' thinking and behaviours (i.e., cognitive style), and an upper-level of stored metacognitive beliefs and plans for processing (i.e., meta-system) that guide processing towards a self-relevant goal.

The low-level of processing can give rise to negative automatic thoughts (e.g., "What if I fail?"). For most people, though a negative thought may lead to some further conceptual processing, this ruminative thinking activity is terminated, and other goals are re-instated. However, the S-REF model proposes that for some people a particularly unhelpful style of self-regulation, termed the cognitive attentional syndrome (CAS), will ensue in response to trigger by a negative thought. The CAS consists of repetitive self-focused thinking in the form of worry and rumination, attentional focus on sources of threat, and unhelpful ways of coping (e.g., avoidance, thought suppression, distraction). This style of self-regulation has the effect of maintaining an individual's sense of threat and distress.

Figure 3.3. Self-Regulatory Executive Function (S-REF) model (Wells, 2009).

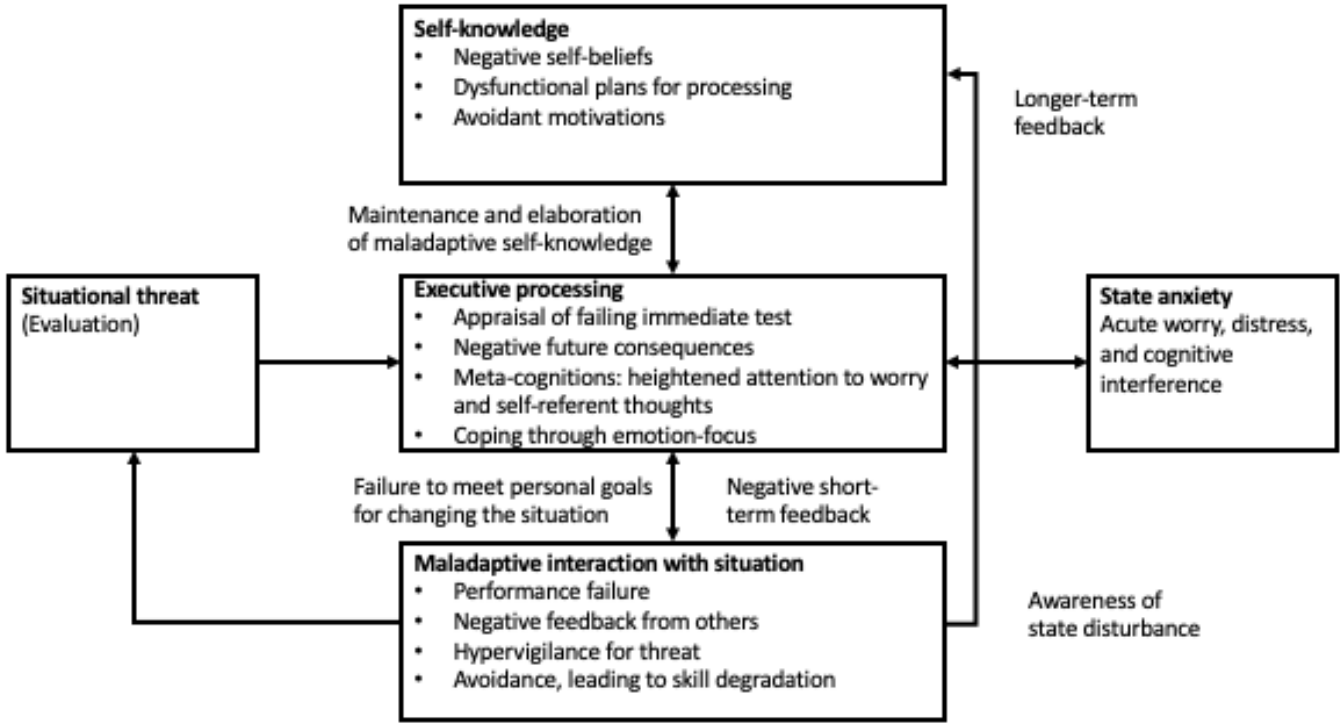


The CAS primarily arises from metacognitive beliefs and knowledge. Metacognition is a broad term, encompassing knowledge, monitoring and regulation of cognitive activity (Flavell, 1979; Nelson & Narens, 1990; Fleur, Bredeweg, & van den Bos, 2021). A broad range of metacognitive beliefs and processes have been implicated across psychological disorders (Sun et al., 2017), including beliefs about: (i) the usefulness of worrying (e.g., “Worrying helps me cope”), (ii) the uncontrollability and danger of worry (e.g., “When I start worrying, I cannot stop”), (iii) one’s memory or ‘cognitive confidence’ (e.g., “I do not trust my memory”), (iv) the need to control thoughts (e.g., “I should be in control of my thoughts all of the time”), and (v) cognitive self-consciousness (e.g., “I pay close attention to the way my minds works”). Negative beliefs about the uncontrollability and danger of worry are considered central to the S-REF model, as they have the effect of locking individuals into the CAS, and thereby perpetuating distress (Wells, 2009).

The S-REF therefore looks beyond beliefs in the cognitive domain, to beliefs and processes in the metacognitive domain that are involved the top-down regulation of thinking and behaviour. This approach implies that modifying maladaptive metacognitive beliefs will remove the CAS, and therefore eliminate sustained distress.

Applying the S-REF model to test anxiety, the CAS may manifest itself as worry about the consequences of failing, active monitoring – of thoughts, emotions, and the environments – for signs of threat (e.g., 'scanning' their body for signs of anxiety or noticing failures in memory), and maladaptive coping efforts (e.g., attempting to suppress thoughts about failure). Manifestation and perpetuation of the CAS will result in test anxiety. Zeidner and Matthews (2005) sketched out an example of how the S-REF model may manifest itself in test anxiety (see Figure 3.4 below).

Figure 3.4. Example S-REF model in test anxiety (Zeidner & Matthews, 2005).



3.3.1 Empirical evidence supporting the S-REF model

There is substantive empirical support for the central tenants of the S-REF model. Metacognitive beliefs are highly correlated with worry in non-clinical (e.g., O'Carroll & Fisher, 2013), mental (Sun et al., 2017), and physical health populations (e.g., Cook et al., 2015). A large corpus of studies has demonstrated that the relationship between emotional disorder and metacognitive beliefs is mediated by the CAS (e.g., Cook et al., 2015; Huntley & Fisher, 2016; Papageorgiou & Wells, 2003). Moreover, the strength of metacognitive beliefs moderates the degree of emotional distress experienced (Bailey & Wells, 2015, 2016). Changes in metacognitive beliefs have been shown to precede changes in distress levels during metacognitive therapy that is based upon the S-REF model (Hoffart et al., 2018).

The excessive tendency to focus on self-relevant threats is a feature of the CAS. Attentional bias is broadly defined as the disproportionate allocation of attentional resource to the selective processing of threat. A great many studies have demonstrated attentional bias experimentally (see reviews by Bar Haim et al., 2007; Cisler & Koster, 2010). High-test-anxious students demonstrate attentional bias towards threatening test-related words (Putwain et al., 2011; Putwain et al., 2020). Attentional bias has largely been considered automatic in nature (Bar Haim et al., 2007; Cisler & Koster, 2010). In contrast, however, the S-REF proposes that attentional bias primarily reflects a voluntary strategy (but one which may become to feel automatic over time). There are several lines of supporting evidence for this contention. Studies of priming effects show that prior presentation of self-relevant material increases interference in both clinical and non-clinical individuals, and over time periods associated with voluntary

processing (see review by Abado, Aue, Okon-Singer, 2020). Connectionist simulation modelling examining automatic versus strategic processing found support for the strategic threat-monitoring (Matthews & Harley, 1996). Metacognitive beliefs, and in particular 'negative beliefs about the uncontrollability and danger of worry', are associated with attentional bias in individuals with health anxiety (Kaur, Butuw, Thewes, 2011).

Coping strategies such as avoidance and self-medication (e.g., alcohol) are generally accepted to exacerbate and perpetuate psychological distress. Many coping strategies are metacognitive in nature are represent voluntary and strategic attempts to deal with unwanted thoughts (Wells, 2009). For example, individuals may attempt to suppress an unwanted negative thought. However, thought suppression can result in rebounding and more negative thoughts (Wegner et al., 1987; Purdon et al., 2005). The lack of effectiveness of strategies such as thought suppression and worry can lead the individual to feel they cannot control their thinking and even give raise to metacognitive judgements that thinking is harmful (Wells, 2009).

Though there is substantive support for the S-REF model as applied to emotional disorders (Sun et al., 2017; Wells, 2000, 2009), only three studies have examined the model as applied to test anxiety in university students (Herzer, Wendt, & Hamm, 2015; Matthews, Hillyard, & Campbell, 1999; Spada et al., 2006). Metacognitive beliefs are related to maladaptive coping in examinations (Matthews et al., 1999), mediate the relationship between test anxiety and a surface approach to studying (Spada et al., 2006), and can help discriminate between students requesting clinical help for their test anxiety and non-clinical controls (Herzer et al. 2015). Four other studies examine metacognitive beliefs

among university students in test anxiety-related constructs, with metacognitive beliefs associated with evaluation anxiety (Rogaten, Moneta, & Spada, 2017; Spada et al., 2012), performance test anxiety (O'Carroll & Fisher, 2012), and state anxiety measured three weeks prior to the students' examination period commencing (Spada, Georgiou, & Wells, 2014). The S-REF model has also been examined as applied to test anxiety in non-university student samples. Greater endorsement in beliefs about self-control were indirectly linked with better examination performance via worry, while cognitive reappraisal was indirectly linked to worse examination via worry, in a sample of high-school students (Putwain, 2019).

Overall, there is substantive empirical support for the S-REF model as applied to worry, anxiety, and emotional disorders, but only preliminary support for its application to test anxiety.

3.4 Summary

In summary, both the IU and S-REF models both propose that repetitive negative thinking in the form of worry, attention to threat, and maladaptive coping result in sustained emotional distress and psychological disorder. A key of point of difference is the IU model proposes beliefs about uncertainty (i.e., *what* individuals think about uncertainty, and therefore beliefs that are in the cognitive domain) are critical to psychological disorder, whereas the S-REF model proposes that *how* a person thinks – and metacognitive beliefs linked to that thinking style – are of greater importance. There is substantive empirical support for both models as applied to emotional problems and disorders. Only two studies have been conducted comparing the relative utility of intolerance of

uncertainty and metacognitive beliefs to worry, distress, and emotional disorders. Negative beliefs about the uncontrollability and danger of worry (and baseline worry severity), but not IU, predicted worry severity among adult frequent worriers (Fergus & Wheless, 2018). IU and negative beliefs about the uncontrollability and danger of worry were differentially associated with emotional disorder symptoms among university students, with both IU and negative beliefs about worry associated with GAD and SAD, IU associated with OCD and panic disorder, negative beliefs about worry associated the MDD, and neither associated with PTSD (Penney, Rachor, & Deleurme, 2020).

Prior to competitively testing predictions derived from the IU and S-REF models, it is first important to ensure the associated primary questionnaires for assessing metacognitive and IU beliefs, namely the Metacognitions Questionnaire – 30 (MCQ-30; Wells & Cartwright-Hatton, 2004) and the Intolerance of Uncertainty Scale – 12 (IUS-12; Carleton et al., 2007), are valid to use in UK university samples.

Chapter Four

Assessing metacognitive and intolerance of uncertainty beliefs among university students and their relationships with test anxiety: Psychometric properties of the Metacognitions Questionnaire – 30 (MCQ-30) and the Intolerance of Uncertainty Scale – 12 (IUS-12)

4.1 Introduction

The intolerance of uncertainty (IU; Dugas et al., 1998; Hebert & Dugas, 2019) and the Self-Regulatory Executive Function (S-REF; Wells & Matthews, 1994, 1996) models are transdiagnostic psychological models of emotional disorder. There is empirical support for transdiagnostic nature of the IU and S-REF models across anxiety and mood disorders (see Chapter 3). However, currently there is little empirical support for the applicability of these models applied to test anxiety. Prior to examinations of the models in test anxiety it is first important that the primary questionnaires assessing IU and metacognitive beliefs, the central feature of the IU and S-REF models respectively, are valid for use among UK university student samples. The Metacognitions Questionnaire – 30 (MCQ-30; Wells & Cartwright-Hatton, 2004) and Intolerance of Uncertainty Scale – 12 (IUS-12; *(*IUS-12; Carleton, Norton, & Asmundson, 2007) are the primary self-report questionnaires for assessing metacognitive and IU respectively. Both questionnaires are short versions of scales originally developed to assess metacognitive beliefs (i.e., MCQ-65; Cartwright-Hatton & Wells, 1997) and IU (i.e., IUS-27; Freeston et al., 1994). The MCQ-30 and IUS-12 are described below.

4.1.1 Metacognitions Questionnaire – 30 (MCQ-30)

The Metacognitions Questionnaire-30 (MCQ-30; Wells & Cartwright-Hatton, 2004) is the primary self-report measure of metacognitive beliefs. Initial evaluations of the psychometric properties of the MCQ-30 found a stable five-factor structure, while concurrent validity was demonstrated between its subscales and measures of anxiety (Spada, Mohiyeddini, & Wells, 2008; Wells

& Cartwright-Hatton, 2004). The five metacognitive belief domains of the MCQ-30 are: (1) 'positive beliefs about worry', which measures the extent to which a person believes worry and other aspects of the CAS are useful (e.g., "Worrying helps me cope"); (2) 'negative beliefs about uncontrollability and danger of worry' that assess to the extent to which a person believes their worry is controllable or dangerous to their mental state (e.g., "My worrying thoughts persist, no matter how I try and stop them"); (3) 'cognitive confidence', which assess confidence in memory (e.g., "My memory can mislead me at times"); (4) 'need to control thoughts', which assess negative beliefs about the consequences of not controlling one's thoughts (e.g., "I should be in control of my thoughts all the time"), and; (5) 'cognitive self-consciousness' that assesses the extent to which a person is aware of and monitors their thoughts (e.g., "I am constantly aware of my thinking"). The five-factor structure of MCQ-30 has subsequently been replicated in epilepsy (Fisher, Cook, & Noble, 2016), Obsessive-Compulsive Disorder (Grøtte et al., 2016), and breast cancer (Cook et al., 2014) populations. Only one study has used the MCQ-30 in test anxiety (Herzer, Wendt, & Hamm, 2015) and although the MCQ-30 has been used in university student samples (e.g., Huntley & Fisher, 2016), no study has examined its factor structure when administered under high stakes conditions (i.e., summative examinations), and furthermore no study examined the relationships between MCQ-30 subscales and test anxiety.

4.1.2 Intolerance of Uncertainty Scale – 12 (IUS-12)

The IUS-12 assesses individual differences in IU (Carleton, Norton, & Asmundson, 2007). The IUS-12 was primarily created to address the inconsistent

factor structure of the original 27-item IUS, in addition to removing some redundant and Generalized Anxiety Disorder (GAD) specific items in order to improve generalizability of use of the instrument as a measure of IU (Carleton et al., 2007). Initial psychometric investigation revealed the IUS-12 to have a stable two-factor structure, excellent internal consistency, and be highly correlated with worry and anxiety (Carleton, Norton, & Asmundson, 2007). The two IUS-12 factors were labeled 'Prospective IU' and 'Inhibitory IU'. 'Prospective IU' refers to the desire for predictability and the propensity to actively seek out information to help reduce uncertainty, while 'Inhibitory IU' refers to reticence in the face of uncertainty (Carleton et al., 2007). 'Prospective IU' and 'Inhibitory IU' were highly correlated (r = .73), which led Carleton and colleagues to suggest that either total score or subscales could be computed (Carleton, Norton, & Asmundson, 2007). However, highly correlated factors does not provide empirical justification for use of a total score (Reise, Moore, & Haviland, 2010). Thus, although the two-factor structure of the IUS-12 was initially replicated in student, clinical, and community samples (Carleton, Collimore, & Asmundson, 2010; Fergus & Wu, 2013), subsequent research applying more sophisticated bifactor measurement models found the IUS-12 to have a strong and reliable general IU factor that explains the majority of common variance, leading to conclusions that the IUS-12 is essentially a unidimensional instrument and that total scores should be used (Bottesi, et al., 2019; Hale et al., 2016; Saulnier, et al., 2019; Shihata, McEvoy, & Mullan, 2018). This is consistent with theoretical conceptualizations of IU that views it as a unidimensional construct (Dugas et al., 1998; Burton & Dugas, 2018). As yet, no study has examined the factor

structure of the IUS-12 in UK student samples or investigated the relationships between IU and test anxiety.

4.1.3 Aims of this study

In summary, the MCQ-30 and IUS-12 have been validated in many different clinical and non-clinical populations (e.g., Fisher, Cook, & Noble, 2016; Fergus & Wu, 2013). However, the psychometric properties of the MCQ-30 and IUS-12 among UK university students and the relationships of metacognitive beliefs and IU with trait and state test anxiety are yet to be investigated. The *Standards for Educational and Psychological Testing* (AERA, APA, NCME; 2014) state that evidence to support validity is required for each new usage of an instrument or assessment (i.e., in new contexts or populations of interest). In this context, validity refers to the evidence that supports the meaningful interpretation of MCQ-30 and IUS-12 scores (Downing, 2003). Evidence to support validity of usage can be derived from investigations of internal structure and relations to other measures or constructs (AERA, APA, NCME; 2014).

This study has two primary aims: (i) to examine if the previously published factor structures of the MCQ-30 and IUS-12 are valid in UK university student samples, and (ii) to examine the concurrent associations between IU and metacognitive beliefs and test anxiety. To ensure robustness and generalizability of the findings, factor structures and associations with test anxiety are investigated in both trait test anxiety that is assessed during term time, and state test anxiety, assessed immediately prior to summative examinations. The MCQ-30 factor structure has been found to be invariant across mental and physical

health samples thus far (e.g., Cook et al., 2014; Fisher et al., 2016; Grøtte et al., 2016), so it is hypothesized that the published five-factor structure will hold here. Given that IU was conceptualized as a unitary construct and recent empirical data supports that contention (Bottesi, et al., 2019; Hale et al., 2016; Saulnier et al., 2019; Shihata, McEvoy, & Mullan, 2018), it is hypothesized that a bifactor model with a strong general factor will best fit the data. As the aim is to examine concurrent associations between MCQ-30 and IUS-12 subscales and test anxiety, no model is specified in advance. However, it is hypothesized that of the metacognitive belief domains, 'negative beliefs about uncontrollability and danger of worry' will have the largest associations with worry and emotionality dimensions of test anxiety, as these beliefs are consistently associated with emotional disorder (e.g., Roussis & Wells, 2006; Wells, 2005; Sun et al., 2017). It is also hypothesized IU will be significantly associated with both state and trait test anxiety.

4.2 Method

4.2.1 Participants and procedure

Participants were undergraduate students from the University of Liverpool. Data were collected by convenience sampling from two different cohorts. The first cohort ('trait test anxiety') consisted of undergraduate students enrolled across university degree programs who completed the IUS-12, MCQ-30, and a trait measure of test anxiety, the Test Anxiety Inventory (TAI; Spielberger, 1980), online during term time. The second cohort ('state test anxiety') were undergraduate medical students who completed paper copies of the MCQ-30, IUS-12 and a state measure of anxiety, the State-Trait Inventory of Cognitive and

Somatic Anxiety (STICSA-State; Ree et al., 2008), approximately 30 minutes before their Objective Structured Clinical Examinations (OSCEs). OSCEs are used in medical training, where students perform a series of clinical tasks while being observed and evaluated by examiners (Harden, 1988). Participants could be entered into a prize draw for Amazon vouchers (first prize of £100, four second prizes of £25). This research was approved by the University's Ethics Committee. Informed consent was obtained from all participants.

4.2.2 Measures

4.2.2.1 Metacognitions Questionnaire–30 (MCQ-30; Wells & Cartwright-Hatton, 2004)

The MCQ-30 has 30 items assessing metacognitive beliefs. The MCQ-30 has five subscales: (i) 'positive beliefs about worry' (e.g., "Worrying helps me cope"); (ii) 'negative beliefs about uncontrollability and danger of worry' (e.g., "When I start worrying I cannot stop"); (iii) 'cognitive confidence' (e.g., "I do not trust my memory"); (iv) 'need to control thoughts' (e.g., "If I could not control my thoughts, I would not be able to function"); and (v) 'cognitive self-consciousness' ("I constantly examine my thoughts"). Items are scored on a 4-point scale from 1 ("*Do not agree*") to 4 ("*Agree very much*"). Subscale scores can range from 6-to-24, with higher scores indicate greater conviction in metacognitive beliefs. The MCQ-30 has sound psychometric properties, including good internal consistency, test-retest reliability, concurrent validity with measures of anxiety and depression, and convergent validity with the MCQ-65 (Spada et al., 2008; Wells & Cartwright-Hatton, 2004).

4.2.2.2 Intolerance of Uncertainty Scale – 12 (IUS-12; Carleton et al., 2007)

The IUS-12 consists of 12 items assessing IU. It has two subscales: (i) 'Prospective IU' (e.g., "One should always look ahead so as to avoid surprises") and (ii) 'Inhibitory IU' (e.g., "When it's time to act, uncertainty paralyses me"). Items are scored on a 5-point scale from 1 ("*Not at all characteristic of me*") to 5 ("*Very Characteristic of me*"). Total and subscale scores can be computed. Total scores can range from 12-to-60, and subscales score can range from 7-to-35 for 'Prospective IU' and 5-to-25 for 'Inhibitory IU', with higher scores indicating greater conviction in IU beliefs. The IUS-12 has sound psychometric properties, including good internal consistency, test-retest reliability, concurrent validity with measures of anxiety and depression, and convergent validity with the IUS-27 (Carleton et al., 2007; Roma & Hope, 2017).

4.2.2.3 Test Anxiety Inventory (TAI; Spielberger, 1980)

The TAI consists of 20 items assessing a student's typical reactions to examinations (i.e., trait test anxiety). It has two subscales: (i) Worry (e.g., "Thinking about my grade on a course interferes with my work on tests"), and (ii) Emotionality (e.g., "I feel confident and relaxed while taking tests" [reverse scored]). Items are scored on a 4-point scale from 1 ("*Almost never*") to 4 ("*Almost always*"). Total or subscales scores can be computed. Total scores range from 20-to-80. Subscale scores range from 8-to-32. Higher scores indicating greater test anxiety. The TAI has sound psychometric properties (e.g., Everson, Millsap, & Rodriguez, 1991; Spielberger, 1980).

4.2.2.4 State-Trait Inventory for Cognitive and Somatic Anxiety – State Subscale (STICSA-S; Ree et al., 2008)

The STICSA-S consists of 21 items assessing an individual's state anxiety. It has two subscales: (i) S-Cognitive Anxiety (e.g., "I think the worst will happen"), and (ii) S-Somatic Anxiety (e.g., "My breathing is fast and shallow"). Items are scored on a 4-point scale from 1 ("*Not at all*") to 4 ("*Very much so*"). Subscale scores range from 10-40 for Cognitive Anxiety and 11-44 for Somatic Anxiety, with higher scores indicating greater state anxiety. The STICSA-S has good psychometric properties, and the scale has previously been used to measure state anxiety in examination contexts (Gros et al., 2007; Ree et al., 2007). For the purpose of this study, S-Cognitive is referred to as S-Worry and S-Somatic is referred to as S-Emotionality.

4.2.3 Data analytic strategy

4.2.3.1 Examination of the factor structures of the MCQ-30 and IUS-12

Confirmatory Factor Analysis (CFA) examined if the established factor structures of the MCQ-30 and IUS-12 fitted the data. Fit of alternative models was explored using Exploratory Factor Analysis (EFA), where models up to and including five factors were tested for the MCQ-30, and up to a bifactor model with two factors for the IUS-12. Analyses used the weighted least squares estimator (WLSMV) recommended for ordinal data. In the correlated-trait models of two-factors or model, factors are assumed to be inter-correlated and therefore an oblique rotation (i.e., Geomin) determined optimal factor loadings. Adequacy of fit for both CFA and EFA models was assessed using three indices of fit, comprising one incremental and two absolute 'misfit' indices. The

incremental index was the Comparative Fit Index (CFI), where for values ≥ .95 indicate adequate fit (Hu & Bentler, 1999). The two absolute 'misfit' indices were the Root Mean Square Error of Approximation (RMSEA), where values < .05 indicate good fit and values between 0.5 – 0.8 indicate adequate fit (Browne & Cudeck, 1992), and the Standardized Root Mean Square Residual (SRMR) where values of SRMR less than .08 are indicative of good fit (DiStefano, Liu, Jiang, & Shi, 2018; Hu & Bentler, 1999). Inter-correlations amongst factors were examined and the internal consistency of each factor was measured using Cronbach's alpha.

Given, that IU was conceived as a unitary construct, and recent evaluations of the IUS-12 factor structure support this (e.g., Bottesi et al., 2019), additional statistical indices were conducted, consistent with current best practice (Rodriguez, Reise, & Haviland, 2016), to assess the bifactor model of the IUS-12. A bifactor model can assess the dimensionality of the instrument and the reliability of the general (i.e., IU) and specific factors (i.e., Prospective, Inhibitory). Dimensionality of the IUS-12 within the bifactor model was assessed with the following indices: Explained Common Variance (ECV), Percent of Uncontaminated Correlations (PUC), and assessment of standardized factor loadings for the general and specific factors. ECV and PUC inform whether a bifactor model with a strong general factor should be modelled as unidimensional or multidimensional (general and specific factors) in structural equation modelling (SEM). ECV is the proportion of common variance across items explained by the general factor relative to the specific factors, where values greater than .70 indicates support for a strong general factor (Rodriguez et al., 2016). Additionally, the proportion of common variance explained in each item

by the general factor is examined by item-explained common variance (I-ECV), where values greater than .80 suggest items primarily reflect the general factor (Stucky & Edelen, 2015). PUC is the proportion of covariance terms which reflect variance from the general dimension and values greater than .70 reflect unidimensionality (Rodriguez et al., 2016).

Model-based reliability of total and specific factors was assessed using the following indices: Omega Hierarchical General (ωH) and Specific (ωHS), construct reliability (H), and Factor Determinacy (FD). The coefficient ωH represents the proportion of systematic variance that can attributed to the general factor, while ωHS reflects the proportion of systematic variance explained by specific factors after partitioning out variance attributable to the general factor. If ωH values are greater than .80 then total scores are considered essentially unidimensional (Rodriguez et al., 2016). Coefficient H represents the proportion of variance explained by the latent variable (i.e., a unidimensional IU factor in this case) relative the variance unexplained by the latent variable. High H values (> .80) suggest a well-defined latent variable that will be replicable across studies (Rodriguez et al., 2016). FD represents the correlation between factor scores and factors, and it is recommended that factor scores are only used for FD greater than .90 (Gorsuch, 1983 cf. Rodriguez et al., 2016).

4.2.3.2 Descriptive statistics

Inter-correlations between MCQ-30 and IUS-12 and other study variables were examined. Internal consistency of MCQ-30 and IUS-12 was assessed using Cronbach's alpha. Independent *t*-tests examined gender and year of study differences in IUS-12 and other study variables.

4.2.3.3 Examination of concurrent associations between metacognitive beliefs / IU beliefs, and trait and state test anxiety

Associations between IU/metacognitive beliefs was assessed in trait and state test anxiety datasets by fitting a structural model in which latent variables for the test anxiety worry and emotionality dimensions were first regressed onto MCQ-30 factors, and then regressed onto IUS-12 factor(s). Model fit was assessed using the same CFI, RMSEA, and SRMR fit indices described in section 4.2.3.1.

All analyses were conducted in M*plus* version 8 (Muthén & Muthén, 2008-2017).

4.3 Results

4.3.1 Sample characteristics

The participant response rates were 54% (291 out of 541) and 52% (463 out of 882) for the trait and state sample datasets respectively, based upon the number of students who accessed the study website (for trait sample) or had access to the study information (state sample). Responders were defined as participants who supplied sufficient data to be included in data analyses. There no missing data points in the trait dataset but there were 27 missing data points in the State dataset (< 0.2% of total data points). Little's MCAR test confirmed these data points as missing completely at random ($\chi^2(2583) = 2655.44$, $p = .157$) and expectation-maximization was used to impute values.

Sample characteristics for trait and state datasets are in Table 4.1. In summary, the mean age of participants in both samples was similar, and both

samples were constituted of more females than males, and where the largest ethnic group in both was White British. The demographic composition of the samples was similar to the university (trait dataset) and School of Medicine (state dataset) populations from which they were drawn.

4.3.2 Factor Structure of the MCQ-30 and IUS-12

4.3.2.1 Factor structure of the MCQ-30

CFA of the published MCQ-30 correlated traits five-factor solution found adequate fit of the model to trait data: $\chi 2$ (395) = 874.05 $p < .001$, CFI = .92, RMSEA = .066 (90% CIs .060 - .072), SRMR = 0.08 (see Table 4.2). EFA confirmed that a correlated traits five-factor model provided the best fit to the data: $\chi 2$ (295) = 428.22 $p < .001$, CFI = .98, RMSEA = .040 (90% CIs .032 - .048), SRMR = .033 (see Table F.1 in Appendix F for item loadings) and significantly superior fit to 4-factor or less models (all rescaled χ^2 difference tests were significant). Minor discrepancies were observed between the EFA-derived model and the published five-factor model. Item 3 loaded higher on 'negative beliefs about uncontrollability and danger of worry' (factor 1) than its own factor of 'cognitive self-consciousness' (factor 4), and item 27 did not load highly on any factor.

Table 4.1. Characteristics of trait (*n* = 291) and state (*n* = 463) samples.

Demographic	**Domains**	**Trait**	**State**
Age	Mean (SD)	20.78 (3.43)	21.87 (2.52)
		n (% of total *N*)	n (% of total *N*)
Gender	Female	205 (70.4%)	259 (56%)
	Male	83 (28.5%)	204 (44%)
Year of study	1	121 (41.6%)	-
	2	89 (30.6%)	200 (43%)
	3	52 (17.9%)	144 (31%)
	4	29 (10.0%)	119 (26%)
Ethnicity	White (British, Irish, other)	228 (78.4%)	313 (68%)
	Chinese	24 (8.2%)	15 (3%)
	Indian subcontinent (Indian, Pakistani, other)	12 (4.1%)	92 (20%)
	Black (British, African, Caribbean, other)	10 (3.4%)	12 (3%)
	Mixed race	10 (3.4%)	13 (3%)
	Other	5 (1.7%)	14 (3%)

Missing data: (trait) Age, *n* = 2; Gender, *n* = 3; Year of Study, *n* = 0; Ethnicity, *n* = 1; (state) Age, *n* = 4; Gender, *n* = 0; Year of Study, *n* = 0; Ethnicity, *n* = 4.

Table 4.2. MCQ-30 published scale structure and Geomin rotated factor loadings from CFA correlated factors model of trait (n = 280) and state (n = 463) test anxiety datasets.

		Correlated factors model - trait					Correlated factors model - state				
MCQ-30 published scale structure and items		**F1**	**F2**	**F3**	**F4**	**F5**	**F1**	**F2**	**F3**	**F4**	**F5**
Subscale: Positive beliefs about worry											
MCQ-1	Worrying helps me to avoid problems in future		**0.72**					**0.73**			
MCQ-7	I need to worry in order to remain organized		**0.82**					**0.80**			
MCQ-10	Worrying helps me get things sorted out in my mind		**0.87**					**0.82**			
MCQ-19	Worrying helps me cope		**0.83**					**0.78**			
MCQ-23	Worrying helps me solve problems		**0.82**					**0.80**			
MCQ-28	I need to worry in order to work well		**0.85**					**0.76**			
Subscale: Negative beliefs about worry											
MCQ-2	My worrying is dangerous for me	**0.67**					**0.68**				
MCQ-4	I could make myself sick with worrying	**0.78**					**0.74**				
MCQ-9	My worrying thoughts persist, no matter how I try to stop them	**0.85**					**0.86**				

MCQ-11	I cannot ignore my worrying thoughts	**0.83**		**0.85**	
MCQ-15	My worrying could make me go mad	**0.81**		**0.81**	
MCQ-21	When I start worrying, I cannot stop	**0.85**		**0.86**	
Subscale: Cognitive Confidence					
MCQ-8	I have little confidence in my memory for words and names		**0.83**		**0.84**
MCQ-14	My memory can mislead me at times		**0.64**		**0.69**
MCQ-17	I have a poor memory		**0.79**		**0.81**
MCQ-24	I have little confidence in my memory for places		**0.76**		**0.82**
MCQ-26	I do not trust my memory		**0.85**		**0.91**
MCQ-29	I have little confidence in my memory for actions		**0.88**		**0.83**
Subscale: Need to control thoughts					
MCQ-6	If I did not control a worrying thought, and then it happened, it would be my fault			**0.71**	**0.73**
MCQ-13	I should be in control of my thoughts all the time			**0.60**	**0.51**
MCQ-20	Not being able to control my thoughts is a sign of weakness			**0.76**	**0.68**

MCQ-22	I will be punished for not controlling certain thoughts	**0.77**	**0.78**
MCQ-25	It is bad to think certain thoughts	**0.61**	**0.64**
MCQ-27	If I could not control my thoughts, I would not be able to function	**0.44**	0.39
Subscale: Cognitive self-consciousness			
MCQ-3	I think a lot about my thoughts	**0.82**	**0.77**
MCQ-5	I am aware of the way my mind works when I am working through a problem	**0.51**	0.33
MCQ-12	I monitor my thoughts	**0.60**	**0.63**
MCQ-16	I am constantly aware of my thinking	**0.86**	**0.86**
MCQ-18	I pay close attention to the way my mind works	**0.74**	**0.71**
MCQ-30	I constantly examine my thoughts	**0.85**	**0.85**

Notes.

F1 = ‘Negative beliefs about worry’, F2 = ‘Positive beliefs about worry’, F3 = ‘Cognitive confidence’, F4 = ‘Cognitive self-consciousness’, F5 = ‘Need for control over thoughts’; Bold = loading ≥ 0.40; Underline = highest loading where item loads ≥ 0.40 on more than one factor.

For the state dataset, CFA of the MCQ-30 five-factor model found adequate model fit: $\chi 2$ (395) = 992.79 $p < .001$, CFI = .95, RMSEA = .057 (90% CIs .053 - .062), SRMR = 0.08 (see Table 4.2). EFA confirmed that a five-factor model again provided the best fit to the data: $\chi 2$ (295) = 625.27 $p < .001$, CFI = .97, RMSEA = .049 (90% CIs .044 - .055), WRMR = .034 (see Table F.1 in Appendix F for item loadings) and significantly superior fit to 4-factor or less models (all rescaled χ^2 difference tests were significant). All items loaded on their expected factors except items 3, 6, 13, and 27. Item 3 loaded highly on its own factor of 'cognitive self-consciousness' (factor 4) but also loaded highly on 'negative beliefs about uncontrollability and danger of worry' (factor 1). Items 6, 13, and 27 did not highly on any other factor.

4.3.2.1 Factor Structure of the IUS-12

CFA of unidimensional, correlated factors, and bifactor measurement models were first examined for trait (Table 4.3) and state test anxiety (Table 4.4).

Model fit statistics for IUS-12 in trait test anxiety dataset revealed all measurement models fit data adequately. Rescaled χ^2 difference tests found the bifactor model fit the data significantly better than the correlated factors model (all rescaled tests χ^2 were significant).

Assessment of the bifactor model in the trait test anxiety sample found a strong general factor, indicated by higher standardized loadings for all general factor items compared to specific factors and the general factor explained 82% of the common variance (ECV) compared to just 8% and 10% explained by 'Prospective IU' and 'Inhibitory IU' specific factors respectively. The mean I-ECV value was .81 (range .65 to .99), with 50% of IUS-12 items having I-ECV

greater than .80, indicating that these items are stronger indicators of the general factor than their specific factors. The PUC value of .53 indicated the general factor accounted for approximately half of item correlations. Bifactor reliability indices support the construct reliability of the general factor (ωH = .87, H = .92, FD = .94).

EFA confirmed that a bi-factor model provided the best fit to the trait test anxiety data: $\chi 2$ (42) = 126.34 p < .001, CFI = .98, RMSEA = .079 (90% CIs .071 - .100), SRMR = .034 (Table F.2 in Appendix F shows item loadings and statistics). Assessment of the bifactor model found a strong general factor, indicated by higher standardized loadings for all general factor items compared

Table 4.3. Standardized factor loadings and confirmatory factor analysis goodness-of-fit indices for measurement models of the Intolerance of Uncertainty Scale – 12 in *trait* test anxiety sample.

	Unidimensional	Correlated factors model		Bifactor model		
Item	**model**	**Inhibitory**	**Prospective**	**General**	**Inhibitory**	**Prospective**
3. Uncertainty keeps me from living a full life.	.33	.35		.39	-.27	
6. When it is time to act, uncertainty paralyzes me.	.75	.79		.68	.49	
7. When I am uncertain, I can't function very well.	.75	.80		.70	.44	
10. The smallest doubt can stop me from acting.	.73	.77		.69	.35	
12. I must get away from all uncertain situations.	.76	.81		.75	.24	
1. Unforeseen events upset me greatly.	.79		.81	.78		.19
2. It frustrates me not having all the information I need.	.78		.80	.72		.41
4. One should always look ahead so as to avoid surprises.	.62		.63	.57		.32
5. A small unforeseen event can spoil everything, even with the best of planning.	.76		.78	.78		.06
8. I always want to know what the future has in store for me.	.77		.81	.76		.27

9. I can't stand being taken by surprise.	.78		.80	.73		.33
11. I should be able to organize everything in advance.	.68		.70	.64		.29
Model fit statistics						
χ^2 *(df)*	253 (54)	172 (53)			126 (42)	
CFI	.95	.97			.98	
RMSEA [90% CI]	.11 (.10 - .13)	.09 (.07 - .10)			.08 (.07-.10)	
WRMR	.05	.04			.03	
Bifactor model ancillary statistics						
% Explained Common Variance (ECV)	-	-	-	0.82	0.08	0.10
% Omega hierarchical	-	-	-	0.87	0.11	0.11
Construct reliability (*H*)	-	-	-	0.92	0.13	0.45
Factor Determinacy (*FD*)	-	-	-	0.94	0.62	0.74

Notes.

CFI = Comparative Fit Index; RMSEA = Root mean square error of approximation; WRMR = Weighted Root Mean square Residual.

Table 4.4. Standardized factor loadings and confirmatory factor analysis goodness-of-fit indices for measurement models of the Intolerance of Uncertainty Scale – 12 in *state* test anxiety sample.

	Unidimensional	**Correlated factors model**		**Bifactor model**		
Item	**model**	**Inhibitory**	**Prospective**	**General**	**Inhibitory**	**Prospective**
3. Uncertainty keeps me from living a full life.	.65	.64		.70	.17	
6. When it is time to act, uncertainty paralyzes me.	.79	.77		.70	.49	
7. When I am uncertain, I can't function very well.	.80	.79		.69	.60	
10. The smallest doubt can stop me from acting.	.78	.77		.71	.28	
12. I must get away from all uncertain situations.	.67	.66		.80	.10	
1. Unforeseen events upset me greatly.	.30		.29	.77		-.08
2. It frustrates me not having all the information I need.	.72		.71	.62		.20
4. One should always look ahead so as to avoid surprises.	.70		.70	.60		.45

5. A small unforeseen event can spoil everything, even with the best of planning.	.67		.67	.81		.01
8. I always want to know what the future has in store for me.	.80		.79	.70		.48
9. I can't stand being taken by surprise.	.77		.76	.79		.20
11. I should be able to organize everything in advance.	.73		.72	.61		.48
Model fit statistics						
χ^2 (df)	465 (54)	451 (60)			168 (42)	
CFI	.93	.93			.98	
RMSEA [90% CI]	.13 (.12 - .14)	.13 (.12 - .14)			.08 (.07-.09)	
WRMR	.06	.06			.03	
Bifactor model ancillary statistics						
% Explained Common variance (ECV)	-	-	-	0.81	0.10	0.10
% Omega hierarchical	-	-	-	0.88	0.10	0.16

Construct reliability (*H*)	-	-	-	0.93	0.48	0.50
Factor Determinacy (*FD*)	-	-	-	0.95	0.76	0.81

Notes.

CFI = Comparative Fit Index; RMSEA = Root mean square error of approximation; WRMR = Weighted Root Mean square Residual.

to specific factors and the general factor explained 82% of the common variance (ECV), compared to just 8% and 10% explained by ‘Prospective IU’ and ‘Inhibitory IU’ specific factors respectively. The mean I-ECV value was .82 (range .66 to .99), with 58% of IUS-12 items having I-ECV greater than .80, indicating that these items are stronger indicators of the general factor than their specific factors. The PUC value of .53 indicated the general factor accounted for approximately half of item correlations. Bifactor reliability indices support the construct reliability of the general factor (ωH = .87, H = .92, FD = .94).

CFA model fit statistics for IUS-12 in state test anxiety dataset also found all measurement models fit data adequately. Rescaled χ^2 difference tests found the bifactor model fit the data significantly better than the correlated factors model (all rescaled χ^2tests were significant). Assessment of the bifactor model found a strong general factor, indicated by higher standardized loadings for all general factor items compared to specific factors and the general factor explained 80% of the common variance (ECV) compared to just 10% and 10% explained by ‘Prospective IU’ and ‘Inhibitory IU’ factors respectively. The mean I-ECV value was .82 (range .56 to .99), with 58% of IUS-12 items having I-ECV greater than .80, indicating that these items are stronger indicators of the general factor than their specific factors. The PUC value of .53 indicated the general factor accounted for approximately half of item correlations. Bifactor reliability indices support the construct reliability of the general factor (ωH = .88, H = .93, FD = .95).

EFA model fit statistics for IUS-12 in the state test anxiety dataset confirmed the bifactor model fit the data best ($\chi 2$ (33) = 127.27 p < .001, CFI = .99, RMSEA = .071 (90% CIs .064 - .085), SRMR = .033 (Table F.3 in

Appendix F shows item loadings and statistics). Assessment of the bifactor model found a strong general factor, indicated by higher standardized loadings for all general factor items compared to specific factors and the general factor explained 86% of the common variance (ECV) compared to just 10% and 4% explained by 'Prospective IU' and 'Inhibitory IU' factors respectively. The mean I-ECV value was .87 (range .71 to .99), with 66% of IUS-12 items having I-ECV greater than .80, indicating that these items are stronger indicators of the general factor than their specific factors. The PUC value of .55 indicated the general factor accounted for approximately half of item correlations. Bifactor reliability indices support the construct reliability of the general factor (ωH = .94, H = .94, FD = .97).

Overall bifactor indices of the IUS-12 in both trait and state test anxiety samples show that specific factors, despite explaining a small proportion of variance, do not possess sufficient reliable variance to enable interpretation, whereas the general factor does have construct reliability. Evidence here supports deriving only total IUS-12 scores.

4.3.3. Descriptive statistics

Means and standard deviations of the MCQ-30 subscales and intercorrelations between them are presented in Table 4.5 (derived from CFA analyses). The majority of intercorrelations were significant, and mostly in the medium-to-large effect size range based upon Cohen's (1992) taxonomy, where .01 is small, .03 is medium, and .05 and above is large.

Table 4.5. Descriptive data and correlations between Metacognitions Questionnaire-30 and test anxiety measures (Trait = TAI; State = STICSA) subscales

	POS	NEG	CC	NC	CSC	TA-W	TA-E	M (SD)	Cronbach's α
Trait									
POS	-	.22***	.16**	.29***	.31***	.13*	.22***	11.99 (4.55)	.88
NEG		-	.34***	.57***	.42***	.54***	.51***	14.84 (5.09)	.87
CC			-	.31***	.07	.41***	.29***	11.88 (4.72)	.87
NC				-	.38***	.40***	.31***	12.51 (3.99)	.74
CSC					-	.20**	.19**	15.94 (4.26)	.84
TA-W						-	.73***	18.53 (6.36)	.90
TA-E							-	22.64 (6.21)	.90
State									
POS	-	.31***	.39***	.46***	.35***	.37***	.28***	11.08 (3.64)	.86
NEG		-	.55***	.51***	.49***	.69***	.62***	12.31 (4.55)	.88
CC			-	.42***	.33***	.59***	.45***	11.32 (4.28)	.88
NC				-	.55***	.43***	.29***	10.53 (3.21)	.70
CSC					-	.40***	.31***	13.71 (4.05)	.82
TA-W						-	.71***	20.72 (6.90)	.90
TA-E							-	23.29 (6.53)	.88

Notes.

MCQ-30 subscales: POS = Positive beliefs about worry; NEG = Negative beliefs about worry; CC = Cognitive confidence; NC = Need to control thoughts; CSC = Cognitive self-consciousness; TA-W = Test Anxiety – Worry; TA-E = Test Anxiety – Emotionality; *** $p < .001$, ** $p < .01$, * $p < .05$.

Table 4.6. Descriptive data and Pearson's r correlations between IUS-12 and test anxiety (Trait = TAI; State = STICSA) subscales.

	IU	**TA-W**	**TA-E**	***M* (*SD*)**	**Cronbach's α**
Trait					
IU	-	.44***	.40***	33.23 (10.56)	.91
TA-W		-	.74***	18.37 (6.42)	.91
TA-E			-	22.37 (6.29)	.90
State					
IU	-	.55***	.43***	28.71 (9.41)	.91
TA-W		-	.71***	20.72 (6.90)	.90
TA-E			-	23.29 (6.53)	.88

Notes.

IU = Intolerance of Uncertainty Scale – 12 total score; TA-W = Test Anxiety – Worry; TA-E = Test Anxiety – Emotionality; M = mean, SD = standard deviation. * $p < .05$, ** $p < .01$, *** $p < .001$.

Internal consistency was good across MCQ-30 subscales with Cronbach's alphas ranging between .82 – .88, except for the 'Need to control thoughts' subscale which had acceptable internal consistency of .70 – .74.

Means and standard deviations of the IUS-12 total score and intercorrelations between them are presented in Table 4.6 (derived from CFA analyses). IU were significantly positively correlated with trait and state test anxiety (Pearson's *r* from .34 to .57). Internal consistency of IUS-12 was excellent (Cronbach's alphas of .91) across trait and state datasets.

Differences in scores due to gender were examined. Given multiple independent *t*-tests are conducted, a Bonferroni correction was applied. Here, the alpha of .05 was divided by 8, the number of tests conducted, to yield a Bonferroni corrected alpha of .001 (rounded up).

In the 'trait' dataset, gender differences were found in trait test anxiety, with females reporting greater worry (M_{Female} = 19.22 vs. M_{Male} = 16.18; $t(281) = 3.68$, $p < .001$) and emotionality (M_{Female} = 23.64 vs M_{Male} = 19.11; $t(281) = 5.80$, $p < .001$) but no gender differences in IU were found. No differences were found in IU, test anxiety worry and emotionality based on students' year of study. Age did not correlate with IUS-12 or any TAI, or STICSA subscale.

In the 'state' dataset, gender differences were found in state test anxiety and IU, with females reporting greater worry (M_{Female} = 21.84 vs. M_{Male} = 19.30; $t(461) = 4.00$, $p < .001$), emotionality (M_{Female} = 24.67 vs. M_{Male} = 21.55; $t(461) = 5.23$, $p < .001$), MCQ-NEG (M_{Female} = 13.06 vs. M_{Male} = 11.35; $t(461) = 4.10$, $p < .001$), MCQ-CC (M_{Female} = 11.90 vs. M_{Male} = 10.58; $t(461) = 3.33$, $p = .001$) and IU (M_{Female} = 29.90 vs. M_{Male} = 27.20; $t(461) = 3.09$, $p = .002$). With regard to differences in test anxiety and IU scores due to year of study, no significant differences were found, except for a significant difference in state test anxiety worry scores ($F(2, 460) = 3.71$, $p = .025$), with Year 2 students reporting significantly less worry than Year 3 students ($M_{difference}$ = -2.03).

4.3.4 Concurrent associations between metacognitive beliefs / IU beliefs and test anxiety

Prior to examining the relationships between IU and metacognitive beliefs and test anxiety in the 'trait' and 'state' datasets, measurement models of

the TAI and STICSA-S were examined, with correlated two-factor models for the TAI ($\chi 2$ (103) = 268.80 p < .001, CFI = .98, RMSEA = .08 (90% CIs .06 - .09), SRMR = .04) and STICSA ($\chi 2$ (188) = 647.15 p < .001, CFI = .95, RMSEA = .07 (90% CIs .07 - .08), SRMR = .05) having acceptable fit.

4.3.4.1 Concurrent associations between metacognitive beliefs and trait and state test anxiety

The relationships between metacognitive beliefs and concurrent test anxiety worry and emotionality dimensions are shown in Figure 4.1 for the trait dataset and Figure 4.2 for the state dataset. Gender was included as a covariate of worry and emotionality in both 'trait' and 'state' structural models, as gender differences in test anxiety are well known (Hembree, 1988; Seipp, 1991; von der Embse et al., 2018). 'Year of study' was also included as a covariate of worry in the 'state' model, as a significant difference between year groups was found suggesting that past examination experience has induced more worry, and so including for this variable will permit examination of this conjectured relationship. Again, for simplicity, relationships between gender and year of study and study variables are not shown in Figures 4.3 and 4.4.

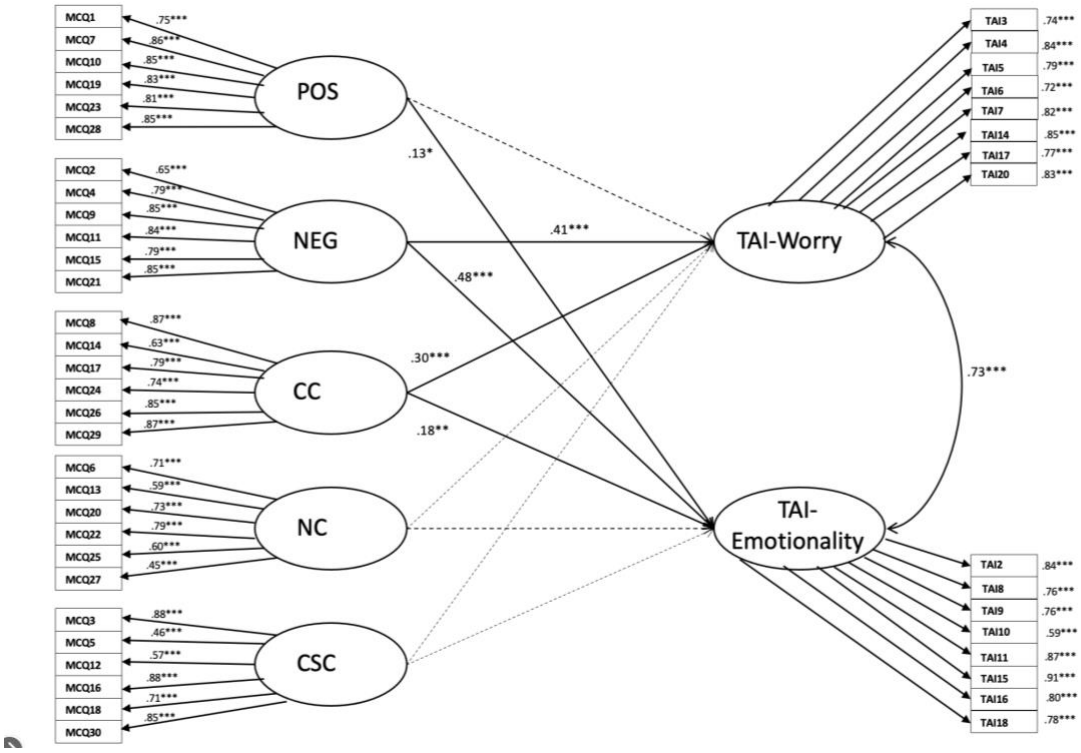
Fit indices indicate acceptable model fit for both trait: $\chi 2$ (1057) = 1685.27 p < .001, CFI = .94, RMSEA = .046 (90% CIs .042 - .050), SRMR = 0.07), and state datasets: $\chi 2$ (1302) = 2337.33 p < .001, CFI = .95, RMSEA = .041 (90% CIs .039 - .044), SRMR = 0.06. As hypothesized, 'negative beliefs about uncontrollability and danger of worry' has the strongest association with both worry and emotionality dimensions of test anxiety across both conditions. However, slightly different relationships were observed between models. For the

trait model, only ‘negative beliefs about uncontrollability and danger of worry’ and ‘cognitive confidence’ were significantly associated with worry. ‘Negative beliefs about uncontrollability and danger of worry’, ‘positive beliefs about worry’, and ‘cognitive confidence’ were all significantly associated with emotionality. Neither ‘need to control thoughts’ nor ‘cognitive self-consciousness’ were associated with worry or emotionality. In the state model, ‘negative beliefs about uncontrollability and danger of worry’, ‘positive beliefs about worry’, and ‘cognitive confidence’ were all significantly associated with both worry and emotionality. ‘Need for control over thoughts’ was also significantly associated with emotionality but a negative relationship was observed. No significant relationships between ‘Cognitive self-consciousness’ and worry and emotionality were found. Finally, there was a stronger relationship between worry and emotionality dimensions in the trait model than in the state model.

4.3.4.2 Concurrent associations between IU and trait and state test anxiety

The relationships between IU and test anxiety worry and emotionality are presented in Figure 4.3 for the trait dataset and Figure 4.4 for the state dataset. Again, gender was included as a covariate of worry and emotionality in both ‘trait’ and ‘state’ structural models, and ‘Year of study’ was also included as a covariate of worry in the ‘state’ model. Again, for simplicity, relationships between gender and year of study and study variables are not shown in Figures 4.3 and 4.4.

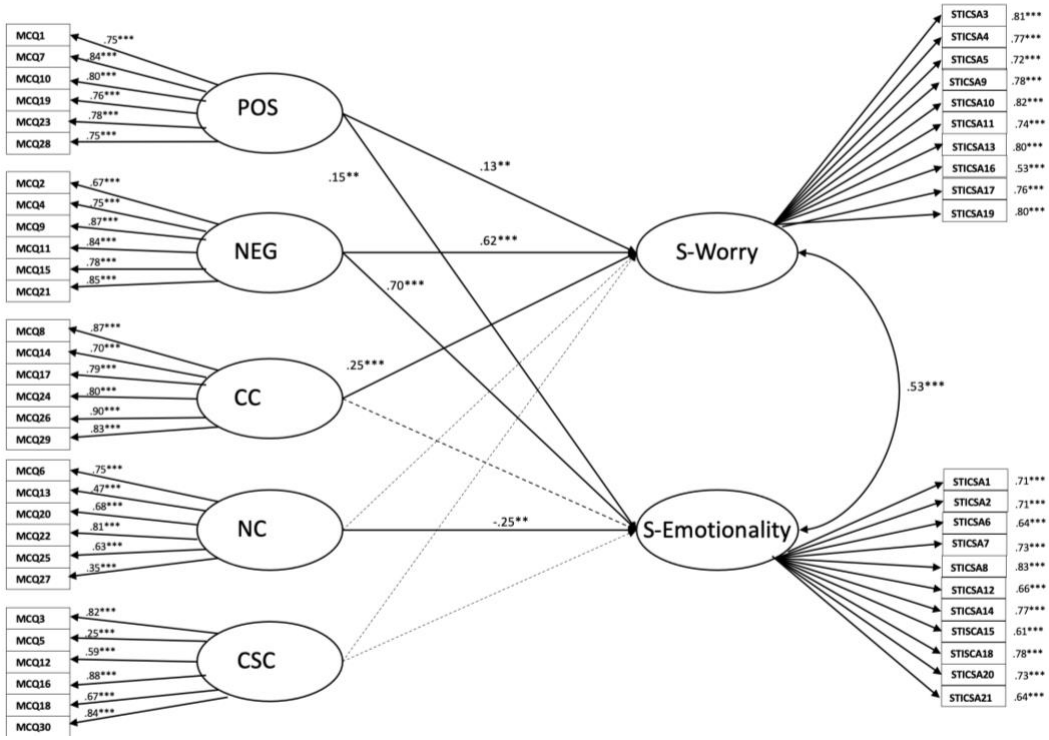
Figure 4.1. Structural equation modeling of the relationships between latent factors of the MCQ and dimensions of (trait) TA.



Notes.

Ellipses indicate latent factors, rectangles indicate observed variables. POS = ‘Positive beliefs about worry’; NEG = ‘Negative beliefs about worry’; CC = ‘Cognitive confidence’; NC = ‘Need to control thoughts’; CSC = ‘Cognitive self-consciousness’; TA-W = Test Anxiety Inventory – Worry; TA-E = Test Anxiety Inventory - Emotionality. Figures show standardized path coefficients. Dotted lines indicate non-significant relationships. * $p < .05$, ** $p < .01$, *** $p < .001$.

Figure 4.2. Structural equation modeling of the relationships between latent factors of the MCQ and dimensions of (state) TA.



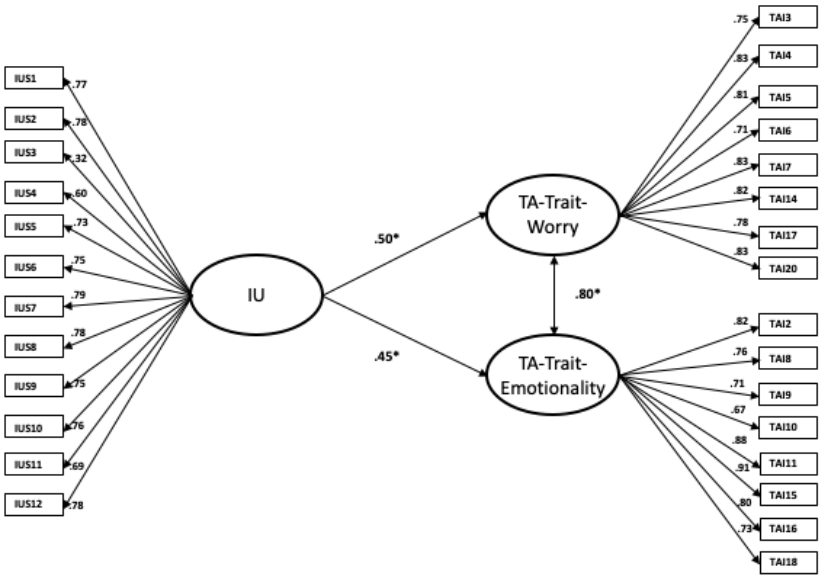
Notes.

Ellipses indicate latent factors, rectangles indicate observed variables. POS = ‘Positive beliefs about worry’; NEG = ‘Negative beliefs about worry’; CC = ‘Cognitive confidence’; NC = ‘Need to control thoughts’; CSC = ‘Cognitive self-consciousness’; S-Worry = State (Test Anxiety) – Worry (as measured by STICSA-S Cognitive subscale); S-Emotionality = State (Test Anxiety) – Emotionality (as measured by STICSA-S Somatic subscale); STISCA-S = State-Trait Inventory for Cognitive and Somatic Anxiety – State. Figures show standardized path coefficients. Dotted lines indicate non-significant relationships. * $p < .05$, ** $p < .01$, *** $p < .001$.

Fit indices indicated acceptable model fit in the ‘trait’ dataset: $\chi 2$ (373) = 647.82 $p < .001$, CFI = .97, RMSEA = .051 (90% CIs .044 - .057), SRMR = .10. IU was significantly and positively associated with test anxiety worry and emotionality dimensions. Gender was not significantly associated with worry or emotionality.

In the ‘state’ dataset, the structural model also had acceptable fit: $\chi 2$ (554) = 1363.20 $p < .001$, CFI = .95, RMSEA = .056 (90% CIs .052 - .060), SRMR = 0.06. IU was significantly and positively associated with test anxiety worry and emotionality dimensions, with a stronger association with worry exhibited. Gender was significantly associated with worry (-.10, $p < .05$) and emotionality (-.18, $p < .05$), and ‘year of study’ was significantly and positively associated with worry (-.08, $p < .05$).

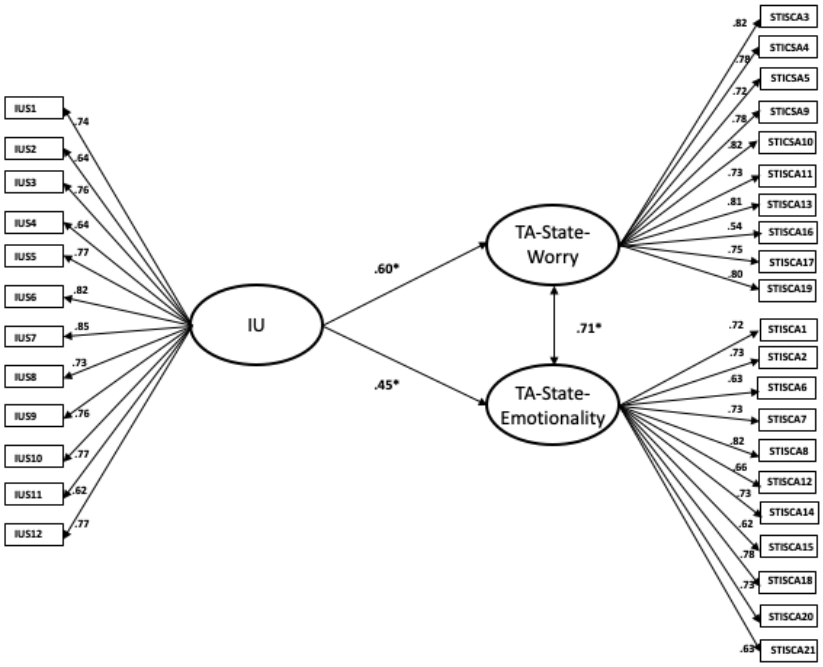
Figure 4.3. Structural equation modeling of the relationships between latent factors of the IUS-12 and dimensions of (trait) test anxiety.



Notes.

Ellipses indicate latent factors, rectangles indicate observed variables. IU = Intolerance of Uncertainty Scale – 12; TA = Test Anxiety (Test Anxiety Inventory). Figures show standardized path coefficients. Dotted lines indicate non-significant relationships. * $p < .05$, ** $p < .01$, *** $p < .001$.

Figure 4.4. Structural equation modeling of the relationships between latent factors of the IUS-12 and dimensions of (state) test anxiety.



Notes.

Ellipses indicate latent factors, rectangles indicate observed variables. IU = Intolerance of Uncertainty Scale – 12; TA = Test Anxiety (as measured by STICSA-S); STISCA-S = State-Trait Inventory for Cognitive and Somatic Anxiety – State. Figures show standardized path coefficients. Dotted lines indicate non-significant relationships. * $p < .05$, ** $p < .01$, *** $p < .001$.

4.4 Discussion

The primary aims of this study were to examine the factor structures of the IUS-12 and MCQ-30 as applied to test anxiety among university students, and to examine the concurrent validity between these constructs and the worry and emotionality dimensions of test anxiety. Firstly, factor analyses supported the previously published five-factor and bifactor model solutions of the MCQ-30 and IUS-12 respectively. Internal consistencies were all in the acceptable-to-excellent range, suggesting MCQ-30 subscale items and IUS-12 items are tapping into their respective latent constructs (e.g., IUS-12 items all measure the intolerance of uncertainty construct). Overall, the MCQ-30 and IUS-12 factor structures appear valid in test anxiety. Secondly, in tests of concurrent validity, structural equation modeling was used to examine relationships between MCQ-30 subscales and IUS-12 and test anxiety worry and emotionality dimensions. Both IU and metacognitive beliefs were significatively and positively associated with both trait and state test anxiety, indicating both may represent important factors in the manifestation of test anxiety.

4.4.1 Metacognitive beliefs and test anxiety

Given the MCQ-30 five-factor structure provided best with the data, the recommendation would be for users of this scale in university samples should compute subscale scores and interpret them accordingly.

With regard to relationships between metacognitive beliefs and test anxiety, ‘negative beliefs about uncontrollability and danger of worry’ had the strongest association with worry and emotionality dimensions of test anxiety, across both datasets. ‘Cognitive confidence’ was also significantly associated

with test anxiety worry and emotionality across both datasets, while ‘positive beliefs about worry’ was associated with both worry and emotionality in the state dataset but only associated with emotionality in the trait dataset. This latter result was surprising as a stronger relationship between ‘positive beliefs about worry’ and the worry dimension of test anxiety, as opposed to emotionality, would be expected, given that ‘positive beliefs about worry’ leads to selection of worry as a coping strategy (Wells & Matthews, 1994, 1996). One possible explanation is that students may be more aware of using worry as a coping strategy in an examination environment than when they reflect on their test anxiety during term time. Finally, a negative relationship was found between ‘need to control thoughts’ and emotionality in the state dataset. This finding is curious and suggests greater beliefs in the need to control thoughts results in lower somatic symptoms. However, this may be a result of the lower internal consistency of the ‘need to control thoughts’ factor and lack of significant loadings for some items. It may also be the case that these beliefs are not viewed negatively by students and that their perception is that being able to distract oneself from unwanted thoughts has a useful short-term effect. Overall, and as predicted ‘negative beliefs about uncontrollability and danger of worry’ was the strongest predictor of test anxiety. ‘Cognitive confidence’ beliefs may also play an important role in test anxiety, where lack of belief in one’s memory may result in increased anxiety about one’s ability to perform well in tests.

4.4.2. Intolerance of uncertainty and test anxiety

Given factor analyses of the IUS-12 found a bifactor model provided best fit to data, with a strong and reliable general IU factor that accounted for the

majority of variance in total IUS-12 scores, then practically the IUS-12 can be considered a unidimensional representation of IU, and users of the instrument in UK university student contexts should compute total scores. Measurement model results from this study are consistent with previous studies examining the factor structure of the IUS-12, which found that a bifactor measurement model provides best fit (Bottesi et al., 2019; Hale et al., 2016; Saulnier et al., 2019; Shihata et al., 2018). In the current study, the strong general IU factor was found to account for much of the reliable variance in the IUS-12 total score and also account for the majority of reliable variance of the specific 'prospective IU' and 'inhibitory IU' subscale scores too. The weak construct reliability of 'inhibitory IU' and 'prospective IU' factors argues against their scoring and interpretation. Thus, the robustness of a bifactor measurement model was confirmed in UK samples for the first time, and the IUS-12 factor structure was demonstrated in both trait and state test anxiety contexts, increasing generalizability of findings, and providing confidence to researchers in this field who wish to use the scale.

Structural models found IU is significantly associated with both trait and state test anxiety, with IU showing marginally stronger associations with the test anxiety worry compared to emotionality dimension. This suggests that IU is an important factor in test anxiety and that reducing IU might effectively alleviate test anxiety. Gender was not associated with trait test anxiety but was significantly associated with state test anxiety, and given females reported higher state anxiety, suggests that females react with greater anxiety in high stakes examination contexts than males. Year of study was associated with the worry dimension of test anxiety suggesting prior examination experience influences the degree of worry in the present examination.

4.4.3 Limitations of this study

There were several limitations to this study. Sample sizes for the trait samples were relatively modest which can reduce reliability of estimates. A cross-sectional design was used whereas prospective designs can elucidate predictive associations between metacognitive beliefs and test anxiety. Finally, students find OSCEs more anxiety provoking than other forms of tests (Brand & Schoonheim-Klein, 2009; Guraya, et al., 2018; Marshall & Jones, 2003; Nicholson & Forrest, 2009), so different patterns of relationships may be found in other test contexts.

4.4.4 Conclusions

Both the IUS-12 and MCQ-30 appear valid measures of IU and metacognitive beliefs respectively among UK university student samples. This permits increased confidence when interpreting results from the next two studies that use these questionnaire measures. Moreover, both IU and metacognitive beliefs were significantly and positively associated with test anxiety, suggesting they play a contributory role in its manifestation. However, it may be that one sets of beliefs plays a more important role in the manifestation of test anxiety than the other. Therefore, the next step is to examine the comparative explanatory utility of both models in explaining test anxiety symptom severity.

Chapter Five

The relative contribution of intolerance of uncertainty and metacognitive beliefs to test anxiety in university students

5.1 Introduction

The transdiagnostic intolerance of uncertainty (IU; Dugas et al., 1998; Hebert & Dugas, 2019) and Self-Regulatory Executive Function (S-REF; Wells & Matthews, 1994, 1996) models can account for worry and somatic anxiety symptoms. A key difference in IU and S-REF models is the set of beliefs considered important to the manifestation of worry and emotional disorder. Beliefs about uncertainty are emphasized in the IU model, while beliefs about cognitions – metacognition – are emphasized in the S-REF model. Although there is empirical support for the IU and S-REF models in anxiety disorders such as Generalized Anxiety Disorder (GAD), there had been scant research done on their applicability to test anxiety (see Chapter 3). Prior to investigation of the IU and S-REF as applied to test anxiety, it was important to establish the validity of the primary questionnaires assessing IU and metacognitive beliefs, the Intolerance of Uncertainty Scale – 12 (IUS-12; Carleton et al., 2007) and Metacognitions Questionnaire – 30 (MCQ-30; Wells & Cartwright-Hatton, 2004) respectively. IUS-12 and MCQ-30 factors structures were replicated in UK university student samples and concurrent validity was demonstrated with both IU and metacognitive beliefs significantly and positively associated with test anxiety (see Chapter 4). However, it may be that one set of beliefs is more central to test anxiety than the other. For example, IU and metacognitive beliefs are differentially associated with emotional disorder symptoms in university symptoms (Penney, Rachor, & Deleurme, 2020). Understanding the contribution that both IU and metacognitive beliefs make to test anxiety will help to develop a better psychological conceptualization of test anxiety, an important first step toward the development of an efficacious intervention.

5.1.1 Aims and hypotheses of this study

The aim of this study is to examine the relative contribution of metacognitive beliefs and IU in test anxiety. Hypotheses are derived from the S-REF model, which posits that metacognitive beliefs are more important than beliefs in the ordinary cognitive domain (i.e., IU) in the genesis and maintenance of emotional disorder (Wells, 2009). For example, metacognitive beliefs have been shown to contribute to generalized anxiety and major depressive disorder symptomology over cognitive beliefs (Myers & Wells, 2005; Huntley & Fisher, 2016; Yilmaz, Gencoz, & Wells 2015). The hypotheses are: (i) both IU and metacognitive beliefs will each make significant unique contributions to test anxiety, (ii) metacognitive beliefs will be able to explain significant additional variance in test anxiety, over-and-above IU, and (iii) that the metacognitive belief domain of 'negative beliefs about the uncontrollability and danger of worry' will make the largest contribution in the final model.

5.2 Method

5.2.1. Participants and Procedure

A cross-sectional design was used to examine the contributions of IU and metacognitive beliefs to students' trait test anxiety (i.e., their typical reactions to tests). A convenience sample of students from the University of Liverpool was recruited. The study was advertised by emails, distributed by departmental administrators, and an announcement on the university student intranet. Participants completed online questionnaires on their test anxiety, IU, and metacognitive beliefs. Questionnaire order was randomized. Informed consent

was obtained from all participants. Participation was voluntary. Those who completed the study had the opportunity to enter a prize draw for Amazon vouchers. The study was granted ethical approval by the University's Ethics Committee.

5.2.2. Measures

5.2.2.1 Test Anxiety Inventory (TAI; Spielberger, 1980).

The TAI has 20-items assessing a student's typical reactions to examinations, i.e., trait test anxiety (e.g., "Thinking about my grade in a course interferes with my work on tests"). Items are scored from 1 ("*Almost never*") to 4 ("*Almost always*"). Total scores are used here and can range from 20-80, with higher scores indicating greater test anxiety. The TAI has good internal consistency, convergent validity, and acceptable test-retest reliability (Spielberger, 1980). Internal consistency of the TAI in this study was excellent with a Cronbach's alpha of .93.

5.2.2.2 Intolerance of Uncertainty Scale – 12 (IUS-12; Carleton, Norton, & Asmundson, 2007).

The IUS-12 has 12 items that assess an individual's IU (e.g., "When I am uncertain, I cannot function very well"). Items are scored from 1 ("*Not at all characteristic of me*") to 5 ("*Entirely characteristic of me*"). Total scores range from 12-60, with higher scores indicating greater IU. The IUS-12 has good internal consistency, convergent validity, and acceptable test-retest reliability (Carleton et al., 2007; Hale et al., 2016) and has support for its validity and use

among UK college student samples (Huntley et al., 2021). Internal consistency of the IUS-12 in this study was excellent with a Cronbach's alpha of .90.

5.2.2.3 Metacognitions Questionnaire – 30 (MCQ-30; Wells & Cartwright-Hatton, 2004).

The MCQ-30 has 30 items that assess metacognitive belief domains. It has five subscales: (i) 'positive beliefs about worry' (e.g., "Worrying helps me to avoid problems in the future"); (ii) 'negative beliefs about uncontrollability and danger of worry' (e.g., "I cannot ignore my worrying thoughts"); (iii) 'cognitive confidence' (e.g., "I do not trust my memory"); (iv) 'need to control thoughts' (e.g., "It is bad to think certain thoughts"); and (v) 'cognitive self-consciousness' (e.g., "I think a lot about my thoughts"). Items are scored from 1 ("*Do not agree*") to 4 ("*Agree very much*"). Subscale scores range from 6-24, with higher scores indicating greater endorsement of metacognitive beliefs. The MCQ-30 has good internal consistency, convergent validity, and acceptable test-retest reliability (Wells & Cartwright-Hatton, 2004), and is valid in UK college student samples (Huntley et al., 2020). Internal consistency of the subscales in this study ranged from acceptable-to-good, (Cronbach's alphas 73 to.89).

5.2.3 Data Analysis Strategy

Variable data were examined for normality via visual inspection of histograms, inspection of skewness and kurtosis coefficients, and Kolmogorov-Smirnov test statistics. Boxplots were inspected to identify univariate outliers. Independent *t*-tests examined for gender differences in independent (IUS-12, MCQ-30) and dependent (TAI) variables. Given multiple independent *t*-tests are

conducted, a Bonferroni correction was applied. Here, the alpha of .05 was divided by 7, the number of tests conducted, to yield a Bonferroni corrected alpha of .001 (rounded up). Homogeneity of variance between male and female groups was examined using Levene's test. Intercorrelations between test anxiety (TAI), IU (IUS-12), and metacognitive beliefs (MCQ-30 subscales) were examined using Pearson's *r*.

Hierarchical multiple linear regression analyses tested the hypotheses that intolerance of uncertainty and metacognitive beliefs would each make significant independent contributions to test anxiety, but that metacognitive beliefs would explain additional variance over-and-above intolerance of uncertainty. The entry method was used enter variables into the regression model, with demographics (age and gender) entered on Step 1, IU entered on Step 2, and metacognitive beliefs entered on Step 3. An additional regression analysis was then performed, in which Steps 2 and 3 were reversed, to examine if IU made an additional contribution in explaining variance in test anxiety severity over-and-above metacognitive beliefs. Robust estimation, using bias corrected and accelerated bootstrapping techniques, which adjust for bias and skewness in the bootstrap distribution, were used (based upon 5,000 samples).

Regression diagnostics were used to assess model assumptions. Multicollinearity was examined by inspection of the correlation matrix, variation inflation factors (VIFs), and tolerance statistics, where correlations greater than 0.8, VIFs greater than 10, and tolerance values less than .2 indicate multicollinearity (Field, 2009). Autocorrelation was assessed using the Durban-Watson statistics, where values close to two indicate the assumption of independent errors is satisfied. The assumption of normally distributed errors

was assessed through inspection of a histogram of standardized residual errors. Heteroscedasticity was assessed via plots of standardized residuals against standardized predicted values. Leverage was evaluated by Mahalanobis distances, where values greater than 25 in a samples greater than 500 participants would be problematic (Field, 2009). Influential cases were assessed using Cook's distances, where values greater than one indicate problematic cases.

All data analyses were conducted with SPSS 25 (IBM, 2017).

5.3 Results

5.3.1 Sample characteristics

Six-hundred and seventy-five students completed the study from a total of 1,389 students who viewed the study website. The demographic composition of the sample is described in Table 5.1. The mean age of the sample was 21.03 years (*SD* = 3.11). The sample was constituted of female (71.4%) and 'White British, Irish, or other' (81.8%) majorities.

Table 5.1. Demographic characteristics of sample (*n* = 675).

Demographic	Domains	Statistics
Age (years)	Mean (SD)	21.03 (3.11)
		n (% of total *N*)
Gender	Female	482 (71.4%)
	Male	188 (27.9%)
Year of study	1	231 (34.2%)
	2	212 (31.4%)
	3	153 (22.7%)
	4	77 (11.4%)
Ethnicity	White (British, Irish, other)	552 (81.8%)
	Chinese or Asian	40 (5.9%)
	Asian subcontinent (Indian, Pakistani, other)	33 (4.9%)
	Black (British, African, Caribbean, other)	16 (2.4%)
	Mixed race	26 (3.9%)
	Arab	4 (0.6%)
	Other	2 (0.3%)

Notes.

Missing data for: Age, *n* = 2; Gender, *n* = 5; Year of Study, *n* = 2; Ethnicity, *n* = 2.

5.3.2 Descriptive statistics

Data screening showed all variables were normally distributed, except ‘positive beliefs about worry’ and ‘cognitive confidence’, which were slightly positive skewed. Box plots revealed no univariate outliers.

Descriptive statistics and inter-correlations between trait test anxiety, IU, and metacognitive belief scores are shown in Table 5.2. Independent *t*-tests found significant differences in scores between genders, with females scoring higher than males on test anxiety and ‘negative beliefs about uncontrollability and danger of worry’, but males scoring higher than females on ‘cognitive self-consciousness’. Levene’s test of homogeneity of variances found significant differences in variances between males and females on the test anxiety (TAI; $F(668) = 5.30$, $p = .022$), IU (IUS-12; $F(668) = 3.848$, $p = .050$) and ‘need to control thoughts’ (MCQ-30-NC; $F(668) = 4.40$, $p = .036$), though the ratio between variances was less than two, and therefore it is safe to assume homogeneity of variances (Field, 2009).

There were significant positive Pearson’s *r* correlations between all study variables, ranging from .12 to .60. Correlations between participant age and the study variables were examined, finding age was only significantly correlated with ‘cognitive self-consciousness’ ($r = .08$, $p = .034$). Partial correlations found that all relationships between study variables remained positively correlated when controlling for gender. Finally, differences in test anxiety, IU, or metacognitive belief domain scores based on the year of study a participant was in were examined; no significant differences were found.

Table 5.2. Zero-order correlations (upper right quadrant) controlling for gender (lower left quadrant), and descriptive statistics for study variables.

								Female	**Male**	
								(*n* = 482)	**(*n* = 188)**	**Gender differences**
Variable	**1.**	**2.**	**3.**	**4.**	**5.**	**6.**	**7.**	**M (SD)**	**M (SD)**	***t*-test statistics**
1. TAI	-	.47***	.19***	.57***	.35***	.33***	.18***	53.78 (14.71)	45.45 (13.50)	$t = 6.74$, $p < .001$
2. IUS-12	.46***	-	.33***	.60***	.33***	.57***	.34***	33.05 (10.40)	31.54 (9.28)	$t = 1.74$, $p = .083$
3. MCQ-POS	.19***	.33***	-	.28***	.16***	.32***	.31***	12.03 (4.45)	11.81 (4.62)	$t = 0.56$, $p = .754$
4. MCQ-NEG	.56***	.60***	.28***	-	.36***	.56***	.40***	15.20 (5.12)	13.13 (5.05)	$t = 4.71$, $p < .001$
5. MCQ-CC	.36***	.36***	.16***	.36***	-	.37***	.12***	11.99 (4.96)	11.69 (4.56)	$t = 0.73$, $p = .469$
6. MCQ-NC	.36***	.58***	.33***	.58***	.38***	-	.44***	11.98 (3.94)	12.49 (3.49)	$t = -1.58$, $p = .116$
7. MCQ-CSC	.23***	.36***	.31***	.43***	.13***	.44***	-	15.28 (4.15)	16.55 (4.24)	$t = -3.54$, $p < .001$

Note. * $p < .05$, ** $p < .01$, *** $p < .001$, *ns* = non-significant. TAI = Test Anxiety Inventory; IUS-12 = Intolerance of Uncertainty Scale – 12; MCQ = Metacognitions Questionnaire – 30; POS = 'Positive beliefs about worry'; NEG = 'Negative beliefs about worry'; CC = 'Cognitive Confidence'; NC = 'Need to control thoughts'; CSC = 'Cognitive self-consciousness'.

5.3.3 Do IU and metacognitive beliefs contribute to test anxiety?

Hierarchical multiple regression analyses examined if IU and metacognitive beliefs both contribute to test anxiety, and if metacognitive beliefs can explain additional variance over-and-above IU (see Table 5.3 below). On Step 1, the demographic variables of age and gender explained 6% of variance in test anxiety severity. On Step 2, IU explained a further 20% of variance. Finally, on Step 3, metacognitive beliefs added 13% of variance. When Steps 2 and 3 were reversed, metacognitive beliefs explained an additional 32% of variance in test anxiety severity on Step 2, over age and gender, while IU then explained a further 2% of variance on Step 3. The final model accounted for 40% of variance in test anxiety severity (R^2 = .40), with gender, IU, and the metacognitive belief domains of 'negative beliefs about the uncontrollability and worry' and 'cognitive confidence' making significant contributions. Regression diagnostics revealed one multivariate outlier; this case was removed and the regression re-run, with no changes in the pattern of results found. Regression diagnostics also revealed no cases exerted undue influences or leverage, residuals were normally distributed, and no autocorrelation, presence of multicollinearity, or heteroscedasticity was found.

A test of equality of regression coefficients was used to determine which belief domain made the largest contribution to the final regression model. 'negative beliefs about uncontrollability and danger of worry' made a significantly greater unique contribution to test severity than IU (F[1, 672] = 32.45, p = .001) and 'cognitive confidence' (F[1, 672] = 35.59, p = .001). There was no significant difference in the unique contribution of IU and 'cognitive confidence' to test anxiety severity.

Table 5.3. Statistics for each step of the regression, predicting test anxiety (n = 668)

		Model statistics			Parameter estimates		
Step	**Variable**	ΔR^2	ΔF	p	β	b (95% CI BCa)	p
1.		.06	22.60	< .001			
	Age				.00	-0.01 (-0.37, 0.34)	.976
	Gender				-.25	-8.33 (-10.64, -6.06)	< .001
2.		.20	181.74	< .001			
	Age				.02	0.09 (-0.25, 0.41)	.601
	Gender				-.22	-7.36 (-9.37, -5.35)	< .001
	IUS-12				.45	0.66 (0.56, 0.75)	< .001
3.		.13	29.48	< .001			
	Age				-.01	-0.04 (-0.35, 0.26)	.823
	Gender				-.16	-5.12 (-6.98, -3.18)	< .001
	IUS-12				.18	0.27 (0.15, 0.38)	< .001
	MCQ-30-POS				.01	0.03 (-0.20, 0.27)	.773
	MCQ-30-NEG				.42	1.21 (0.95, 1.45)	< .001
	MCQ-30-CC				.15	0.46 (0.26, 0.67)	< .001
	MCQ-30-NC				-.05	-0.19 (-0.53, 0.15)	.255
	MCQ-30-CSC				-.02	-0.09 (-0.35, 0.18)	.526

Notes.

BCa 95% CI = Bias corrected and accelerated 95% confidence intervals; TAI = Test Anxiety Inventory; IUS = Intolerance of Uncertainty Scale – 12; MCQ =Metacognitions Questionnaire – 30; POS = Positive beliefs about worry; NEG = Negative beliefs about the uncontrollability and danger of worry; CC = Cognitive Confidence; NC = Need to control thoughts; CSC = Cognitive Self-consciousness.

5.4 Discussion

This study is the first to examine the relative contributions of IU and metacognitive beliefs to test anxiety, in order to develop a better understanding of test anxiety among university students as a first step towards developing a new, more efficacious intervention than currently exists. Females reported significantly greater test anxiety than males, and both IU and metacognitive beliefs were significantly positively correlated with test anxiety, even when controlling for gender. Gender, IU and the metacognitive belief domains of 'negative beliefs about the uncontrollability and danger of worry' and 'cognitive confidence' all made significant unique contributions to test anxiety. Of the significant independent contributors to test anxiety, 'negative beliefs about the uncontrollability and danger of worry' made the largest unique contribution to the final regression model.

These results provide empirical support for both the IU and S-REF models. However, results specify a greater role for maladaptive metacognitive beliefs in test anxiety. Specifically, after controlling for IU, metacognitive beliefs explained an additional 13% of variance of test anxiety severity, but when the order was reversed, IU was only able to explain an additional 2% in variance after metacognitive beliefs were controlled. Moreover, the final regression model indicated 'negative beliefs about uncontrollability and danger of worry', which lock individuals into cycles of negative thinking, such as worrying about the consequences of failing a test, and particularly important to test anxiety. If individuals believe their worry is uncontrollable, they will not attempt to cease it and reorient towards being focused on the test and will therefore suffer the consequent damage to their test performance. Therefore, modification of

‘negative beliefs about uncontrollability and danger of worry’ would be important in alleviating test anxiety. IU also contributed to test anxiety, which is consistent with previous findings (Study 1, Chapter 4). This indicates that difficulty tolerating the uncertainty of the tests and their consequences is a significant contributory factor to test anxiety and suggests that students need help in formulating and enacting effective coping strategies to deal with uncertainty. Distrust in one’s memory, or ‘cognitive confidence’ was also significantly associated with test anxiety. This is perhaps not surprisingly given that student’s often report ‘mind blanks’ in test situations (Galassi, Frierson, & Sharer, 1981a cf. Zeidner, 1998). Moreover, as anxiety interferes with cognitive performance (including working memory) and test taking (Angelidis, et al., 2019; Maloney, Sattizahn, & Beilock, 2014; Moran, 2016; von der Embse et al., 2018), it may lead highly test anxious individuals to doubt their mental abilities, including their memory. However, the relationship between subjective confidence in one’s memory and actual memory performance is tenuous at best (Kurdi et al., 2017). Interventions for test anxiety should therefore seek to address metacognitive memory judgements. It is possible that modification of both metacognitive beliefs and IU are needed to reduce the severity of test anxiety.

5.4.1 Limitations of this study

This study has two primary limitations. First, although the sample size was adequate, there were a relatively high proportion of females (71%), which may bias findings given that females report greater test anxiety severity than males. Second, and most importantly, though a cross-sectional design is appropriate for elucidating unique contributions of IU and metacognitive beliefs

in typical test anxiety reactions (i.e., trait test anxiety), prospective designs that examine if intolerance and metacognitive beliefs are causal predictors of test anxiety severity in test situations (i.e., ‘state’ test anxiety) are now required.

5.4.2 Conclusions

Overall, both IU and metacognitive beliefs are implicated in test anxiety, and so the IU and S-REF models appear applicable to test anxiety. However, as a cross-sectional design was used in this study, directional inferences on the role of IU and metacognitive beliefs cannot be drawn. Therefore, it is important to investigate if IU and metacognitive beliefs can prospectively predict test anxiety.

Chapter Six

Do intolerance of uncertainty and metacognitive beliefs predict test anxiety in university students, and does test anxiety mediate their relationship to test performance?

6.1 Introduction

The previous study (see Chapter 5) found both intolerance of uncertainty (IU) and the metacognitive beliefs of 'negative beliefs about the uncontrollability and danger of worry' and 'cognitive confidence' were significantly associated with test anxiety, with 'negative beliefs about the uncontrollability and danger of worry' being most strongly associated. However, the study used a cross-sectional design so the directionality of effects could not be drawn; whilst it is hypothesized that IU or metacognitive beliefs result in test anxiety, it may be that test anxiety leads to IU or metacognitive beliefs. This study addresses this limitation by using a prospective research design to examine if IU and metacognitive beliefs predict test anxiety. Another goal of the current study is to explore the relationships between these beliefs, test anxiety, and test performance in an Objective Structured Clinical Examination (OSCE). Objective Structured Clinical Examinations (OSCEs) are the main form of assessment in undergraduate medical and dental education (Newble, 2004). In OSCEs, students demonstrate their clinical skills and underpinning knowledge across a series of tasks, usually while being observed by an examiner (Gormley, 2011). All candidates are presented with the same tasks, under the same conditions, and are assessed using structured marking schemes (Harden & Glesson, 1979; Gormley, 2011). Clinical competencies that may be assessed included: communication and medical history taking (e.g., gathering information from a patient presenting with breathing difficulties), physical examination skills (e.g., performing a respiratory examination), clinical reasoning (e.g., interpreting clinical data, determining likely diagnosis and management plan), and practical and technical skills (e.g., carry out peak flow respiratory function test). Structured assessment in this way

reduces the bias and lack of reliability of conducting assessments on clinical placements (Gormley, 2011).

6.1.1 Aims and hypotheses

The aims of this study are twofold: (i) to examine if IU and metacognitive beliefs prospectively predict test anxiety, and (ii) examine the relationships between IU/metacognitive beliefs, test anxiety, and test performance. The following predictions, derived from the Self-Regulatory Executive Function (S-REF; Wells & Matthews, 1994, 1996) model are made: (1) both IU and metacognitive beliefs will predict test anxiety, (2) metacognitive beliefs will prospectively contribute to test anxiety over-and-above IU, (3) 'negative beliefs about the uncontrollability and danger of worry' will be the strongest predictor of state test anxiety, and (4) test anxiety will significantly mediate the relationship between maladaptive metacognitive beliefs and IU, and test performance. As past research has shown worry is more strongly associated with test performance than emotionality (Hembree, 1988; von der Embse et al., 2018), the impact of both dimensions on test performance was examined.

6.2. Method

6.2.1 Participants and Procedure

Participants were undergraduate medical students at the University of Liverpool. Students were asked to complete self-report questionnaires at two time points, approximately three months apart (*M* = 96 days). At Time 1, during term time, students completed online questionnaires assessing their IU and metacognitive beliefs. At Time 2, students completed paper versions of

questionnaires assessing their state test anxiety, approximately 30 minutes prior to sitting their summative OSCE (Time 3). At the University of Liverpool, students in Years 2 – 4 sit a summative OSCE. Participation was voluntary and students could withdraw at any time. This study was approved by the University of Liverpool's ethics committee. Informed consent was obtained from all participants.

6.2.2 Measures

6.2.2.1 State-Trait Inventory for Cognitive and Somatic Anxiety – State Subscale (STICSA-S; Ree et al., 2008).

The STICSA-S consists of 21 items assessing an individual's state anxiety. It has two subscales: (i) Cognitive State Anxiety (STICSA-Cog), which consists of 10 items (e.g., "I think the worst will happen"), and (ii) Somatic State Anxiety (STICSA-Som), which consists of 11 items (e.g., "My breathing is fast and shallow"). Participants indicate how much they agree with each statement at the time of completion on a 4-point scale from 1 ("not at all") to 4 ("very much so"). Scores can range from 21-84, with higher scores indicating greater state anxiety. Internal consistency of the STICSA-S in this study, as measured by Cronbach's alpha, was .94. The STICSA-S has sound psychometric properties and has previously been used to measure state test anxiety (Grös et al., 2007; Ree et al., 2008).

6.2.2.2 Intolerance of Uncertainty Scale – 12 (IUS-12; Carleton, Norton, & Asmundson, 2007).

The IUS-12 consists of 12 items assessing IU (e.g., "When I am uncertain, I can't function very well"). Participants indicate how much they agree with each statement on a 5-point scale from 1 ("*Not at all characteristic of me*") to 5 ("*Entirely characteristic of me*"). Scores range from 12-60, with higher scores indicating greater IU. The IUS-12 has sound psychometric properties (Carleton, Norton, & Asmundson, 2007; Roma & Hope, 2017; Huntley et al., 2020). Internal consistencies of the IUS-12 subscales in this study, as measured by Cronbach's alpha, ranged from .73 to .85.

6.2.2.3 Metacognitions Questionnaire – 30 (MCQ-30; Wells & Cartwright-Hatton, 2004).

The MCQ-30 consists of 30 items assessing maladaptive metacognitive beliefs and processes. It has five subscales: (i) 'positive beliefs about worry' (e.g., "Worrying helps me cope"); (ii) 'negative beliefs about uncontrollability and danger of worry' (e.g., "When I start worrying, I cannot stop"); (iii) 'cognitive confidence' (e.g., "I have a poor memory"); (iv) 'need to control thoughts' (e.g., "I should be in control of my thoughts all the time"); and (v) 'cognitive self-consciousness' (e.g., "I monitor my thoughts"). Participants indicate how much they generally agree with each statement on a 4-point scale from 1 ("*Do not agree*") to 4 ("*Agree very much*"). Subscale scores range from 6-24, with higher scores indicating greater maladaptive metacognitive beliefs. The MCQ-30 has sound psychometric properties (Wells & Cartwright-Hatton,

2004; Huntley et al., 2020). Internal consistencies of the MCQ-30 subscales in this study, as measured by Cronbach's alpha, ranged from .69 to .89.

6.2.3 Data analysis strategy

Firstly, variable data were screened. Normality was checked by visual inspection of histograms and kurtosis and skewness statistics. Box plots were inspected for outliers.

Differences in scores on dependent and independent variables between genders were explored using independent *t*-tests. Given multiple independent *t*-tests are conducted, a Bonferroni correction was applied. Here, the alpha of .05 was divided by 7, the number of tests conducted, to yield a Bonferroni corrected alpha of .001 (rounded up). Associations between IU, metacognitive beliefs, and state test anxiety were examined using Pearson's *r* correlations.

Hierarchical multiple linear regression analyses were used to test the main hypotheses that Time 1 IU and metacognitive beliefs would predict state test anxiety, and that metacognitive beliefs would explain additional variance in Time 2 state test anxiety, after first controlling for gender and Time 1 IU. Variables were entered in the regression model in the following steps: (1) gender, (2) Time 1 IU (IUS-12), and finally (3) Time 1 metacognitive beliefs. Additionally, steps 2 and 3 were then reversed to see if the IU explained additional variance over-and-above maladaptive metacognitive beliefs. The entry method was used to enter variables into the model. The statistical significance of R^2_{change} was used to assess if additional variance had been explained.

The following regression diagnostics were used to assess the validity of the model (see Field, 2009): (i) Pearson's *r* correlations, variance inflation

factors (VIFs), and tolerance statistics assessed multicollinearity, where correlations greater than 0.8, VIFs greater than 10, and tolerance values less than .2 indicate multicollinearity; (ii) independent of errors was assessed by the Durban-Watson statistics, where values between 1 and 3 indicate the assumption is satisfied; (iii) assumption of normally distributed errors was assessed via visual inspection of the standardized residual errors histogram; (iv) heteroscedasticity was assessed via visual inspection of standardized residual-standardized residual error plot; (v) Mahalanobis were used to assess leverage of cases, with values greater than 15 problematic in samples of 100; and (v) Cook's distances assessed influential cases, with values greater than 1 cause for concern. A test of equality of regression coefficients was used to examine if 'Negative beliefs about the uncontrollability and danger of worry' made a significantly larger contribution to the final model than other statistically significant predictors.

The PROCESS plugin (Hayes, 2018) was used to test the hypotheses that the test anxiety components of worry and somatic symptoms will mediate the relationship between the metacognitive/IU beliefs and OSCE performance. OSCE scores were standardized (M = 50, SD = 15) before analyses. The significance of the indirect relationships between metacognitive beliefs and IU, and OSCE performance via worry or somatic symptoms was assessed by Aroian test statistic (i.e., this statistic tests if the mediator 'carries' the effect from the predictor to the outcome variable). The Aroian test has the advantage over the traditional Sobel mediation test as it does not make the assumption that the product of the standard errors between the predictor-mediator and mediator-outcome variable are vanishingly small (MacKinnon et al., 2002).

Robust estimation, using bias corrected and accelerated bootstrapping techniques, which adjust for bias and skewness in the bootstrap distribution, were used in both regression and mediation analyses (based upon 5,000 samples).

6.3 Results

6.3.1 Sample characteristics

Of the 174 participants who took part in Time 1, 134 (77%) did so again at Time 2. Of those who completed measures across both time points, 95 (71%) were female and 38 (29%) were male, with one participant not responding. The mean age of the sample was 21.12 years (*SD* = 2.33). With regards to ethnicity, 95 (71%) identified as White, 13 (10%) as from the Asian subcontinent, eight (6%) as Chinese, seven (5%) as Black, seven (5%) as dual heritage, three (2%) as being from another ethnic group, and one participant did not respond. The demographic composition of the sample at Time 2 did not differ markedly from that of Time 1 (see Table 6.1).

6.3.2 Descriptive statistics

Data screening revealed all variables were normally distributed except ‘positive beliefs about worry’, ‘negative beliefs about the uncontrollability and danger of worry’, and ‘cognitive confidence’ which were positively skewed. Inspection of box plots revealed no univariate outliers.

Descriptive statistics and correlations are presented in Table 6.2. Independent *t*-tests revealed significant gender differences in scores, with females reporting significantly greater state test anxiety (M_{female} = 49.91 vs. M_{male} = 42.87, $t[131] = 2.81$, $p = .006$,), and ‘negative beliefs about uncontrollability

and danger of worry' (M_{female} = 14.09 vs. M_{male} = 11.37, t[131] = 2.98, p = .003), while males reported significantly greater 'cognitive self-consciousness' (M_{female} = 14.81 vs. M_{male} = 16.26, t[131] = -2.04, p = .044). All correlations were statistically significant except between 'cognitive confidence' and the following variables: 'positive beliefs about worry', and 'cognitive self-consciousness', and also cognitive confidence and IU when controlling for gender. Age did not significantly correlate with any study variable. One-way ANOVAs found no significant differences in study variable scores based on the year of study a participant was in.

Table 6.1. Sample characteristics at Time 1 (*n* = 175) and Time 2 (*n* = 134).

Variable		**Time 1**	**Time 2**
Age:	*M (SD)*	20.98 (2.37)	21.12 (2.33)
Gender:	Female % (n)	72.0% (126)	70.9% (95)
	Male % (n)	26.9% (47)	28.4% (38)
Year of study:	Year 2	54.9% (96)	50.7% (68)
	Year 3	28.6% (50)	28.4% (38)
	Year 4	16.0% (28)	20.1% (27)
Ethnicity:	White (British, Irish, other)	74.3% (130)	70.9% (95)
	Indian subcontinent	9.7% (17)	9.7% (13)
	Chinese or Asian	5.1% (9)	6.0% (8)
	Black (British, other)	4.6% (8)	5.2% (7)
	Other ethnic group	1.7% (3)	2.2% (3)
	Mixed heritage	4.0% (7)	5.2% (7)

Notes.

Missing at Time 1: age, *n* = 2; gender, *n* = 3; year of study, *n* = 1; ethnicity, *n* = 2;

Missing at Time 2: age, *n* = 0; gender, *n* = 1; year of study, *n* = 1; ethnicity, *n* = 1.

Table 6.2. Intercorrelations and descriptive statistics between study variables (*n* = 133).

Variable	**1.**	**2.**	**3.**	**4.**	**5.**	**6.**	**7.**	**Female (n = 95)**	**Male (n = 38)**	**Gender difference**
								M (SD)	**M (SD)**	**(*t*-test)**
1. T2 STICSA-S	-	.33***	.18*	.63***	.20*	.19*	.18*	49.91 (12.54)	42.87 (14.24)	2.81**
2. T1 IUS-12	.33***	-	.38***	.44***	.17*	.51***	.29**	31.86 (9.23)	30.87 (8.97)	0.57
3. T1 MCQ-30-POS	.21*	.39***	-	.30***	.01	.24**	.18*	12.43 (4.35)	13.55 (5.11)	-1.28
4. T1 MCQ-30-NEG	.61***	.44***	.35***	-	.31***	.42***	.31***	14.09 (4.94)	11.37 (4.30)	2.98**
5. T1 MCQ-30-CC	.19*	.17	.02	.30***	-	.29**	.03	11.28 (4.14)	10.63 (4.61)	0.79
6. T1 MCQ-30-NC	.22*	.52***	.23**	.46***	.30**	-	.38***	11.76 (3.40)	12.42 (3.66)	-0.99
7. T1 MCQ-30-CSC	.23**	.31***	.16	.37**	.04	.37***	-	14.81 (3.55)	16.26 (4.12)	-2.04*

Notes.

* $p < .05$, ** $p < .01$, *** $p < .001$; T1 = Time 1; Time 2; STICSA-S = State-Trait Inventory of Cognitive and Somatic Anxiety – State version; MCQ = Metacognitions Questionnaire -30; POS = 'Positive beliefs about worry'; NEG = 'Negative beliefs about uncontrollability and harm of worry'; CC = 'Cognitive Confidence'; NC = 'Need to control thoughts'; CSC = 'Cognitive Self-Consciousness'; IUS-12 = Intolerance of Uncertainty Scale – 12.

6.3.3 Do metacognitive beliefs and intolerance of uncertainty predict (state) test anxiety?

Hierarchical regressions examined the unique contribution of IU and metacognitive beliefs in predicting state test anxiety, as measured by the STICSA-S (see Table 6.3). On Step 1, gender explained 6% of the variance in state test anxiety. On Step 2, IU explained an additional 10% of variance. Finally, on Step 3, maladaptive metacognitive beliefs explained a further 27% of the variance in state test anxiety. When Steps 2 and 3 were reversed, metacognitive beliefs explained an additional 36% of variance on Step 2, while IU did not make a statistically significant contribution on Step 3. In the final regression model, only 'negative beliefs about uncontrollability and danger of worry' (MCQ-NEG) was a significant independent predictor. The final model explained 43% of the variance (R^2 = .43). Regression diagnostics supported the validity of the final model, with no evidence of multicollinearity, autocorrelation, or heteroscedasticity found. Moreover, no cases exerted undue influence or leverage on the model. One multivariate outlier was found; removal of this outlier did not change the pattern of results.

6.3.4 Does test anxiety mediate the relationship between 'negative beliefs about the uncontrollability and danger of worry' and test performance?

In the mediation analyses, only 'negative beliefs about the uncontrollability and danger of worry' was examined as the independent variable, as it was the only significant predictor of state test anxiety. Independent *t*-tests found females reported both greater test anxiety worry ($t(131) = 2.67$, $p = .009$) and somatic symptoms ($t(131) = 2.52$, $p = .013$) than males. Female OSCE

performance did not significantly differ from that of males ($t(131) = 0.84$, $p = .401$).

The mediation model found ‘negative beliefs about the uncontrollability and danger of worry’ significant predicted state test anxiety worry and somatic symptoms (i.e., emotionality), while only worry significant and negatively predicted OSCE performance (see Figure 6.1). Tests of the indirect pathways found a significant indirect pathway between ‘negative beliefs about the uncontrollability and danger of worry’ and OSCE performance via test anxiety worry (Aroian = -3.07, $p = .002$) but not via somatic symptoms (Aroian = 1.19, $p = .235$).

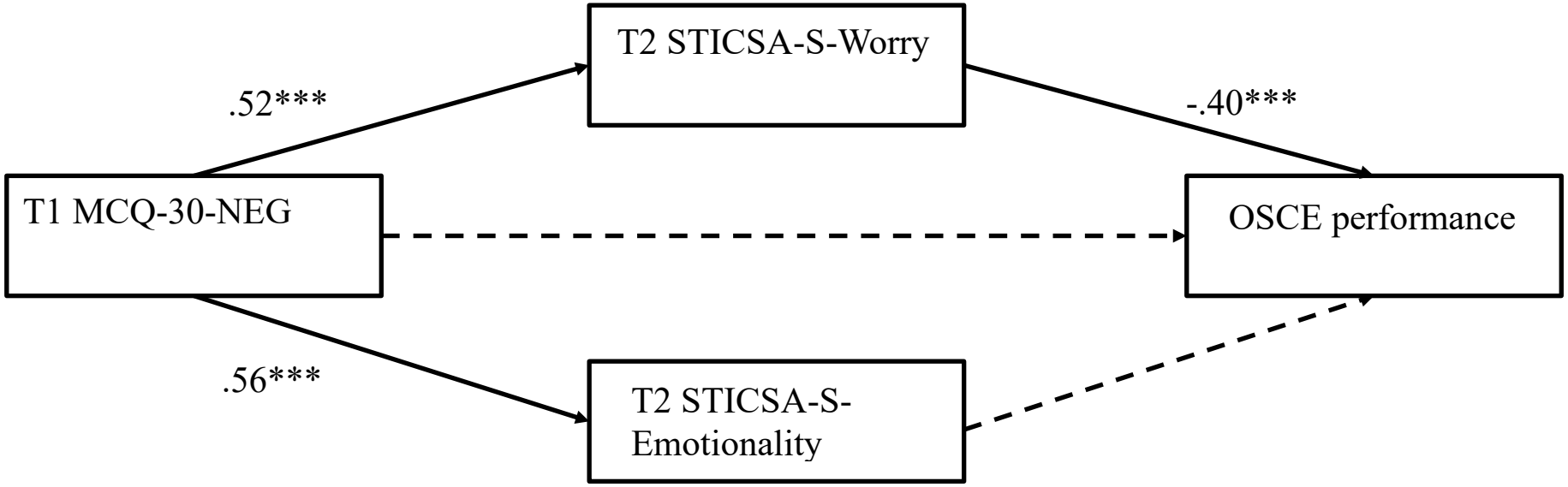
Table 6.3. Hierarchical linear regression, predicting Time 2 state test anxiety (STICSA-S).

Step	Variable	ΔR^2	ΔF	*p*	*b (95% BCa CIs)*	*β*	*p*
1		.06	7.90	.006			
	Gender				-7.04 (-12.11, -1.73)	-.24	.010
2		.10	16.17	< .001			
	Gender				-6.57 (-11.41, -1.63	-.22	.012
	T1 IUS-12				0.47 (0.26, 0.68)	.32	< .001
3		.27	11.59	< .001			
	Gender				-1.92 (-6.77, 3.12)	-.07	.438
	T1 IUS-12				0.20 (-0.08, 0.45)	.14	.154
	T1 MCQ-30-POS				-0.06 (-0.55, 0.39)	-.02	.791
	T1 MCQ-30-NEG				1.67 (1.17, 2.18)	.61	< .001
	T1 MCQ-30-CC				0.09 (-0.33, 0.48)	.03	.669
	T1 MCQ-30-NC				-0.56 (-1.34, 0.22)	-.15	.150
	T1 MCQ-30-CSC				.09 (-0.51, 0.64)	.02	.780

Notes.

STICSA-S = State-Trait Inventory of Cognitive and Somatic Anxiety – State version; MCQ = Metacognitions Questionnaire-30; POS = ‘Positive beliefs about worry’; NEG = ‘Negative beliefs about uncontrollability and harm of worry’; CC = ‘Cognitive Confidence’; NC = ‘Need to control thoughts’; CSC = ‘Cognitive Self-Consciousness’; IUS-12 = Intolerance of Uncertainty Scale – 12.

Figure 6.1. Mediation model, examining if test anxiety (STICSA-S) components (Worry, Emotionality) mediate the relationship between MCQ-30 'Negative beliefs about the uncontrollability and harm of worry' and OSCE performance, whilst controlling for gender.



Notes.

* $p < .05$, ** $p < .01$, *** $p < .001$; T1 = Time 1; T2 = Time 2; MCQ-30 = Metacognitions Questionnaire – 30; NEG = 'Negative beliefs about the uncontrollability and danger of worry"; STICSA-S = State-Trait Inventory of Cognitive and Somatic Anxiety – State subscale; OSCE = Objective Structured Clinical Examination.

6.4 Discussion

This study is the first to examine the causal role of IU and maladaptive metacognitive beliefs in test anxiety, and in turn the relationship with test performance. Both IU and metacognitive beliefs were significantly and positively correlated with state test anxiety. However, metacognitive beliefs explained significant variance (27%) in state test anxiety, after first controlling for gender and IU. When the order of entry was reversed IU did not explain additional variance in state test anxiety after controlling for gender and metacognitive beliefs. In the final regression model, only the metacognitive belief domain of 'negative beliefs about the uncontrollability and danger of worry' was a significant predictor of state test anxiety, indicating it plays a causal role in test anxiety. Finally, worry but not somatic symptoms, was a significant and negatively associated with OSCE performance, and there was a significant indirect pathway from 'negative beliefs about the uncontrollability and danger of worry' and OSCE performance through the worry component of state test anxiety.

6.4.1 Intolerance of uncertainty and metacognitive beliefs as predictors of test anxiety

Findings here show that maladaptive metacognitive beliefs, but not IU, play an important role in the manifestation state test anxiety, and also establish the direction of relationship between metacognitive beliefs and test anxiety. In particular, 'negative beliefs about the uncontrollability and danger of worry' made the largest contribution in predicting distress (i.e., state test anxiety), a finding that consistent with previous investigations of metacognitive beliefs in

mental health populations (Sun, Zhu, & So, 2017). Students with high ‘negative beliefs about the uncontrollability and danger of worry’ will feel they are unable to terminate or suspend their worry, which will have the effect of elevating and maintaining their test anxiety, and therefore interfering with their examination performance.

6.4.2 The relationship between metacognitive beliefs, test anxiety, and test performance

This hypothesis is supported by the findings of the mediational analysis that showed worry led to significantly worse OSCE performance, and that worry was predicted by ‘negative beliefs about the uncontrollability and danger of worry’. No significant relationship between somatic anxiety symptoms (i.e., emotionality) and OSCE performance was found. These findings add to the long list of research that shows that worry is the main feature of test anxiety and one of the key drivers behind poor performance in test situations (Cassady & Johnson, 2002; von der Embse et al., 2018). This study extends previous findings to show that ‘negative beliefs about the uncontrollability and danger of worry’ is a significant driver of worry and therefore should be the target of therapeutic modification. It also first study and the first to examine the impact of test anxiety upon OSCE performance.

6.4.3 Limitations of current study

The current study has several limitations. The sample size is relatively small which may affect the stability of estimates. Sampling bias is indicated given the majority of the sample was female, and females in this study reported

greater state test anxiety than males. Future studies should seek to replicate the findings here, namely worry but not emotionality negatively impacts OSCE performance, in larger samples. Although the central components of both the IU and S-REF models were examined (i.e., IU and metacognitive beliefs respectively), it may be fruitful in future studies to also assess the other components of both models (e.g., assessing aspects of the Cognitive Attentional Syndrome in the S-REF model).

6.4.4. Conclusions

Overall, it appears that metacognitive beliefs, but not IU, are an important determinant of test anxiety in university students, and specifically medical students. Thus, findings here provide support for the S-REF model but not the IU model. Moreover, the worry, but not the emotionality dimension, of test anxiety is associated with poorer OSCE performance.

Chapter Seven:

General Discussion

7.1 Introduction and re-orientation to thesis aims

Test anxiety impairs examination performance and is associated with other mental health problems. It is therefore important to reduce test anxiety. However, the efficacy of different intervention approaches for test-anxious university students was not known. Therefore, the first aim of this thesis was to examine the relative efficacy of interventions for test anxiety compared to control conditions for (i) reducing test anxiety, and (ii) improving test and academic performance.

The second overarching aim was to explore the utility of two transdiagnostic psychological models of emotional disorder, the Intolerance of Uncertainty (IU) and Self-Regulatory Executive Function (S-REF) models, to better understand test anxiety. Each model proposed different beliefs that result in worry and anxiety, with negative beliefs about uncertainty key to the IU model, and metacognitive beliefs central to the S-REF model. Although both models have been extensively evaluated in mental health disorders (Sun et al., 2017; Shihata et al., 2016), little research had investigated their applicability to test anxiety. This thesis therefore represents the first systematic investigation of the comparative utility of the IU and S-REF models to test anxiety. Hypotheses derived from these models are examined in a series of linked cross-sectional (studies 2 and 3) and prospective (study 4) empirical studies. The main findings and implications of these studies are discussed below.

7.2 How efficacious are existing interventions for test anxious university students?

Many interventions for treating test-anxious university student had been evaluated in randomized controlled trials (RCTs). Though two meta-analyses had examined the efficacy of interventions for test anxiety (Hembree, 1988; Ergene, 2003) they had severe methodological limitations. The primary limitation was that both meta-analyses either synthesized data from all educational stages (i.e., from primary/elementary school up to university level) when examining efficacy of different intervention approaches (Hembree, 1988; Ergene, 2003) or synthesized across all intervention approaches for university students (Ergene, 2003). This meant the efficacy of different intervention approaches for test anxious university was unknown. Therefore, an up-to-date meta-analysis that conformed to the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA; Moher et al., 2009; Page et al., 2020) standards for methodological and reporting rigour was required (study 1).

Systematic searches located 44 eligible RCTs with sufficient data to be included in the meta-analyses. Psychological and educational interventions were significantly superior to control conditions for reducing test anxiety at post-treatment and follow-up, and for reducing state test anxiety assessed before examinations. Interventions were also significantly superior to control conditions for improving academic performance at post-treatment but not at follow-up. Overall, interventions achieved moderate effects (g = -0.76) for reducing test anxiety but had weak effects for improving academic performance (g = 0.37). Despite these largely positive findings, the meta-analytic review highlighted five key points of concern that suggests that intervention effects are weaker that

results indicate. First, evidence of publication bias was found, suggesting studies with less promising results have not been published (i.e., the file-drawer effect). Second, inspection of the forest plots often led to the identification of outliers. For example, when examining the efficacy of interventions for reducing test anxiety at post-treatment, three outliers were identified, and removal of these studies led the summary effect size (Hedges' g) reducing substantially, from -0.76 to -0.64. Third, methodological quality assessments found only 16% of items were rated as low risk-of-bias. It is well established that poorer conducted trials produce stronger effect sizes than higher quality trials (Krauss, 2018). Fourth, sample sizes of the included studies were relatively small (means of 15.8 and 17.1 participants in active and control conditions respectively), and smaller samples result in greater bias in outcomes (Borenstein et al., 2009). Fifth, firm conclusions about the longer-term efficacy of interventions cannot be drawn as only 34% studies had follow-up assessments, and the mean follow-up period was just 9.4 weeks. Overall, the quality of studies was poor. It is likely then that effect sizes are inflated, and current interventions achieve moderate effect sizes at best. Therefore, more efficacious interventions are needed, and ultimately larger, higher quality RCTs are required to evaluate these new interventions.

7.3 *Validity of the MCQ-30 and IUS-12 among UK university students*

The meta-analytic review identified the need to develop more efficacious interventions for test anxiety in university students. It was also apparent that none of the included interventions were based on models that could explain the initiation and maintenance of worry, the core feature of test anxiety. Two prominent models of emotional disorder that can explain worry and anxiety are

the IU and S-REF models. Before predictions from these two models could be tested it was important to first establish the validity of the primary measures of IU and metacognitive beliefs, the Intolerance of Uncertainty Scale – 12 (IUS-12) and the Metacognitions Questionnaire – 30 (MCQ-30). The factor structure and concurrent associations with both trait and state test anxiety were examined for both the MCQ-30 and IUS-12 in UK university student samples (Study 2). The published factor structures of both the MCQ-30 and IUS-12 were confirmed. Concurrent associations between IU and metacognitive, and test anxiety were examined using structural equation modeling. The metacognitive beliefs 'positive beliefs about worry', 'negative beliefs about the uncontrollability and danger of worry', and 'cognitive confidence' were significantly and positively associated with both trait and state test anxiety, while 'need to control thoughts' was significantly negatively associated with test anxiety in the state sample only. IU was significantly and positively associated with state and trait test anxiety. Results suggest that both the MCQ-30 and IUS-12 are valid questionnaires to use among UK university student samples.

7.4. Testing theoretical predictions derived from the S-REF model

With support for the validity of the MCQ-30 and IUS-12 established, these questionnaires were subsequently used to test predictions derived from the S-REF model. The S-REF model proposes that beliefs involved in the regulation of cognition (i.e., metacognition) are more important to the manifestation of emotional disorder than beliefs in the cognitive domain (i.e., IU). Four overarching predictions were tested across the two linked cross-sectional (Study 3) and prospective studies (Study 4):

(i) IU and metacognitive beliefs will be associated with test anxiety

(ii) Metacognitive beliefs will explain additional variance in test anxiety over-and-above IU

(iii) The metacognitive belief domain ‘negative beliefs about the uncontrollability and danger of worry’ will be the largest unique predictor in the final regression models

(iv) The relationship between ‘negative beliefs about the uncontrollability and danger of worry’ and test performance will be mediated by test anxiety

Study-specific findings have been discussed in Chapters 5 and 6. Here, findings are collated and discussed with respect to the four overarching predictions stated above.

7.4.1 Current and prospective associations of IU and metacognitive beliefs with test anxiety

Both IU and metacognitive beliefs were hypothesized to be significantly associated with test anxiety. This hypothesis was examined cross-sectionally with trait test anxiety the outcome variable (Study 3, Chapter 5) and prospectively where state test anxiety was the outcome variable (Study 4, Chapter 6).

All IU and metacognitive beliefs were consistently significantly and positively correlated with trait and state test anxiety.

Regression modelling found, as predicted, ‘negative beliefs about uncontrollability and danger of worry’ were most strongly associated with both

trait and state test anxiety. ‘Cognitive confidence’ and IU were found to be significantly associated with trait but not state test anxiety. Neither ‘positive beliefs about worry’, ‘need to control thoughts’ nor ‘cognitive self-consciousness’ were associated with trait or state test anxiety.

Overall, findings are consistent with predictions derived from the S-REF model, which propose that ‘negative beliefs about the uncontrollability and danger of worry’ maintain and exacerbate distress (i.e., test anxiety in this case). Findings also suggest that ‘cognitive confidence’ and IU may be implicated in test anxiety. Students who think their memory is poor may believe they will not be able to recall learnt information under examination conditions, and therefore have greater test anxiety. Likewise, uncertainty about the examination (e.g., format, difficulty) may also lead to greater test anxiety. However, the case for IU and ‘cognitive confidence’ being critical to the manifestation of test anxiety is weaker given that evidence for these beliefs is derived from cross-sectional data only.

7.4.2 Metacognitive beliefs will explain additional variance in test anxiety over-and-above IU

In both Studies 3 and 4, a competitive test of the S-REF and IU models was conducted using hierarchical multiple linear regressions to examine the relative contribution of IU and metacognitive beliefs in explaining variance in test anxiety. Metacognitive beliefs were able to explain a significant amount of additional variance in test anxiety symptom severity, over-and-above IU, age, and gender in both cross-sectional and prospective analyses, findings consistent with theoretical predictions derived from the S-REF. For example, in test

anxiety, a student may have the worry about "I am going to fail" but this thought will naturally decay if it is not engaged with. However, metacognitive beliefs may lead to student to engage in these types of thoughts. Worry may be initiated due believing it to be a useful way of coping with anxiety (i.e., positive beliefs about worry). 'Negative beliefs about the uncontrollability and danger of worry' will then 'lock in' students to worrisome thinking, elevate distress, and use precious mental resources that would be best directed towards the examination.

IU may also play a role in test anxiety. Cross-sectionally, IU explained a significant, albeit small, amount of variance in test anxiety even after controlling for metacognitive beliefs. Uncertainty about examinations may act as an initial trigger for worry. However, prospectively IU did not explain any significant additional variance after controlling for metacognitive beliefs. This suggests that IU may not play a causal role in test anxiety.

Overall, the findings here support predictions of the S-REF model that metacognitive beliefs are more critical to test anxiety than beliefs in the ordinary cognitive domain (i.e., beliefs about uncertainty). Put simply, *how* you think is more important than *what* you think in manifestation of test anxiety.

7.4.3 Negative metacognitive beliefs about the uncontrollability and danger of worry will make the largest unique contribution to test anxiety

'Negative beliefs about the uncontrollability and danger of worry' are considered especially important in the S-REF model, as they perpetuate distress, and bias efforts to try and control thinking, ultimately leading individuals to feel they lack control of their self-regulation (Wells, 2009). These beliefs have been consistently implicated in emotional disorders (Sun et al., 2017). In both final

regression models (Studies 3 and 4), it was predicted that they would be the largest unique predictor of test anxiety. This prediction was supported. These results support the contention that negative beliefs about the uncontrollability and danger of worry appear key in differentiating highly test anxious from their less test anxious peers (Herzer et al., 2015). Although many students may worry, it is only those with high negative beliefs about worry that will struggle to disengage from worry. Overall, these findings support the notion that 'negative beliefs about the uncontrollability and danger of worry' are particularly important in the manifestation and perpetuation of distress. Interventions for test anxious university student should try to reduce the strength of these beliefs if they are to successfully treat test anxiety.

7.4.4 The relationship between 'negative beliefs about the uncontrollability and danger of worry' and examination performance will be mediated by test anxiety

'Negative beliefs about the uncontrollability and danger of worry' was found to be the only significant prospective predictor of test anxiety (Study 4). It was hypothesized that the relationship between examination performance and 'negative beliefs about the uncontrollability and danger of worry' would be mediated by test anxiety. Mediation modelling found there was a significant indirect effect from 'negative beliefs about the uncontrollability and danger of worry' via the worry dimension but not emotionality dimension of test anxiety, to examination performance. Consistent with the wider test anxiety literature, worry was associated with impaired examination performance. Persistent non-task thinking will ultimately interfere with a student's ability to successfully

respond correctly in examinations. OSCEs demand a high cognitive load. For example, students may be asked to calculate correct drug dosages whilst still engaging effectively with the patient, all in the presence of an examiner. Therefore, test anxious students undertaking OSCEs may be particularly at risk of poorer performance compared to more traditional examination formats.

7.5 Limitations of studies within this thesis

Findings here showed that existing interventions for test anxious students have limited efficacy and therefore new interventions are required. The empirical studies here began to illustrate the applicability of two widely used transdiagnostic models of emotional disorders to test anxiety. Overall, results offered greater support for the applicability S-REF model in test anxiety than the IU model, and these findings may inform the development of future interventions to relieve test anxiety. However, there are several limitations that deserve attention.

7.5.1 Limitations of the meta-analysis

This meta-analysis was the first to examine the relative efficacy of different therapeutic approaches (e.g., BT) for treating test anxiety. However, it may be fruitful in the future to examine the relative efficacy of more specific interventions (e.g., systematic desensitization). New statistical techniques such as network meta-analysis have been developed that enable statistical comparisons of intervention outcomes, even if direct comparisons had not been examined in RCTs (Efthimiou et al., 2016). However, such statistical techniques are still developing, as they have been subjected to criticisms that they lead to inflated

and overconfident statements of relative intervention effects (Efthimiou & White, 2019).

The meta-analysis included a substantial number of RCTs (k = 44). However, another 34 RCTs were identified that met eligibility criteria but could not be included due to insufficient available data. This may bias results of this meta-analysis as it is likely inadequately reported RCTs may not have been conducted with the same methodological rigour as better reported RCTs. Furthermore, the loss of data from these studies represents a serious loss of information to the test anxiety literature.

Future RCTs should be conducted and reported to the highest standards to help 'sell' the interventions to educators and wellbeing services. These RCTs should be guided by Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT; Chan et al., 2013) and Consolidated Standards of Reporting Trials (CONSORT; Moher et al., 2010) guidance for developing a trial protocol and reporting trial results respectively. These guidance documents are primarily oriented towards medical clinical trials, so additional factors will need to be accounted for in the planning of psychological oriented interventions, such as the production of a therapeutic manual, and checks on therapist competence and adherence to the intervention protocol (Ost, 2008).

7.5.2 Limitations in empirical studies testing theoretical models

This thesis investigated the key beliefs emphasized by the IU and S-REF models. However, as specified in Chapter 3, both models include other components. Positive worry beliefs, negative problem orientation, cognitive avoidance, and safety behaviours form part of the IU model (Dugas et al., 1998;

Hebert & Dugas, 2019), while the cognitive attentional syndrome (CAS), which consists of perseverative thinking, threat monitoring, and maladaptive coping efforts, is central to the S-REF model. To gain a fuller understanding of both models, future research should and model the relationships between all components, other key variables (e.g., gender), and test anxiety.

Another significant methodological limitation is the cross-sectional nature of studies 2 and 3. This research design does not permit causal inferences to be drawn. Thus, although hypotheses imply causality in one direction (i.e., IU and metacognitive beliefs lead to test anxiety), it may that test anxiety gives rise to IU and metacognitive beliefs. This is unlikely, given that temporal precedence of IU and metacognitive beliefs to other emotional disorders has been demonstrated (e.g., Papageorgiou & Wells, 2009), but it cannot be ruled out. Study 4 addressed this issue by using a prospective research design. Indeed, a different pattern of results emerged with only 'negative beliefs about the uncontrollability and danger of worry' predicting test anxiety in the prospective design (Chapter 6) but gender, IU beliefs, 'negative beliefs about the uncontrollability and danger of worry', and 'cognitive confidence' were all predictors in the cross-sectional design (study 3, Chapter 5).

A further issue is that while trait test anxiety was the outcome in study 3, state test anxiety was the outcome variable in study 4. It may be that 'negative beliefs about the uncontrollability and danger of worry' were only a significant predictor in this sample, but the other variables are all important factors in the general propensity to experience test anxiety (i.e., trait test anxiety). Further prospective and longitudinal research to examine the causal role of IU and metacognitive beliefs is now required.

7.5.3 Sample representativeness

In all datasets there was a predominance of female participants of around 70% of the total samples, representing a respondent bias. It is unclear why this occurred. Perhaps one reason is that females tend to report greater test anxiety than males, and it may be that they hoped to gain understanding of their test anxiety through completion of the study questionnaires. The differential gender response rates may bias statistical estimates given females report greater test anxiety severity than males (Hembre, 1988; Seipp & Schwarzer, 1996). It should be noted that test anxiety study samples do frequently comprise a greater proportion of females (e.g., Pate et al., 2021; von der Embse, Putwain, & Francis, 2021). Additionally, most participants in the studies were of White British heritage. To further generalize results found here larger, more representative samples, and samples from different countries, cultures, and contexts are required.

7.5.4 Examining the relationships between IU and metacognitive beliefs and other putative factors involved in test anxiety and test performance

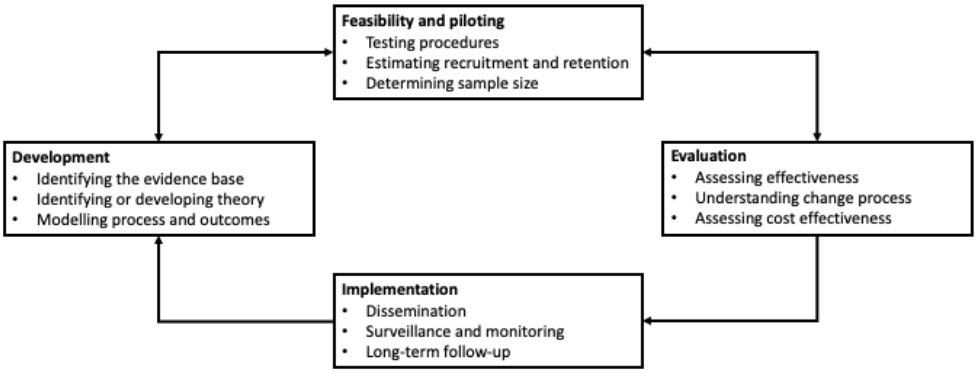
Several other cognitive and personality-related factors have been linked to test anxiety, and is important to examine their relationships with IU and metacognitive beliefs. For example, self-efficacy is significantly positively associated with university academic performance (Richardson et al., 2010). Lower self-efficacy is associated with higher test anxiety (e.g., Yerdelen, McCaffrey, Klassen, 2016), while procrastination is positively associated with test anxiety (Van Eerde, 2003; Steel, 2007). Therefore, examination of the

relationships between maladaptive metacognitive beliefs, self-efficacy, and procrastination, and test anxiety may illuminate if these constructs interact and whether a temporal causal relationship exists between them (e.g., does perseverate thinking mediate the relationships between metacognitive beliefs and procrastination, self-efficacy, and test anxiety?). Overall, it would be expedient to examine a broader range of variables involved in learning and test anxiety in future research whilst examining the IU and S-REF models. A more complete understanding of test anxiety correlates, concomitments, and consequences, would help identify pertinent outcomes for intervention studies.

7.6 Future research development in higher education settings

There are two broad but overlapping areas for future research. The Medical Research Council (MRC; Craig et al., 2008; O'Cathain et al., 2019) offers a framework for the development and evaluation of complex interventions that is considered best practice. The framework advocates a systematic approach to the development of interventions. The framework was presented in Chapter 1 but he re-presented here. The framework consists of four phases (Figure 7.1). Phase 1 involves establishing what is known about current interventions and if new interventions are needed. If new interventions are needed, then theoretical models should be identified, applied, and empirically tested. Phase 2 comprises developing an intervention based upon the theoretical model, examining its feasibility, and conducting pilot tests (e.g., to check on acceptability of the treatment and possible side effects). Phase 3 requires formal evaluation of the intervention, typically via a RCT. Finally, Phase 4 entails the implementation and evaluation of the intervention into routine practice.

Figure 7.1. Medical Research Council Framework for developing and evaluating complex interventions.



7.6.1. Further empirical tests of theoretical models of test anxiety

This thesis has fulfilled the most of aims of Phase 1 (development). That is, the meta-analysis here established what is known about current interventions and identified the need for more efficacious interventions. The applicability of two psychological models to test anxiety were then empirically examined. Findings offered promising support for the S-REF model. However, further investigation S-REF model is still required. Firstly, there is a need to examine whether the CAS mediates the relationship between metacognitive beliefs and test anxiety. Currently, the only all-in-one questionnaire of the CAS is the CAS-1 (Wells, 2009). The CAS-1 consists of non-disorder specific items assesses key aspects of the CAS. Disorder specific versions of the CAS exist for GAD, PTSD, OCD, and MDD (Wells, 2009). It is important to have sufficiently specific measures to accurately examine relationships between psychological constructs in test anxiety, and therefore a test anxiety specific version of the CAS-1 should be developed and validated.

Secondly, and extending from the recommendation above, further prospective, and longitudinal, research is required. The time interval between Time 1 and Time 2 in the prospective study in this thesis was relatively short at approximately three months. Studies with longer follow-up periods, and at least three waves of data collection, would permit explication of the temporal relationships between metacognitive beliefs, CAS processes, and test anxiety.

Thirdly, experimental studies are needed to examine the causal relationships between metacognitive beliefs and test anxiety. Manipulation of metacognitive beliefs and resulting change in distress would provide compelling evidence for the S-REF model. For example, 'cognitive confidence' was an independent predictor of trait test anxiety, suggesting lack of confidence in one's memory may be important factor. Studies that examine the associative and casual relationships between metacognitive beliefs (i.e., cognitive confidence), test anxiety, working memory capacity, worry, and task performances are needed. Working memory capacity of high-test-anxious students is impaired compared to their low-test-anxious peers when doing tasks involving emotional but not neutral cues (e.g., Dutke & Stober, 2001; Shi, Gao, & Zhou, 2014). An intervention focusing on enabling students to make an accurate assessment of their memory abilities could reduce their test anxiety.

7.6.2 Developing interventions to help test anxious university students

Finally, and the goal of these suggested research pathways, is to develop more ways of helping to support students and treating their test anxiety, such as developing interventions that can be ultimately integrated into undergraduate curricula.

Even if new interventions are developed, it can be challenging to embed them within routine practice (e.g., May 2011). Normalisation process theory (NPT), which is routed in sociological theory, can be used to evaluate the likelihood of successfully integrating complex new interventions into existing services and organizational infrastructures (May et al., 2007). NPT focuses on identifying the factors that facilitate or deter the implementation of a new intervention (May et al., 2007), like a test anxiety intervention in a higher education for test anxiety in higher education setting. There are four key concepts in NPT concerning how stakeholders make sense of new interventions, how they can implement them within their practice, the wider work that must be undertaken at strategic and organizational levels, and finally the appraisal and reflections on the implementation of the new intervention. NPT was originally applied in healthcare settings (e.g., May et al., 2011) but has since also been applied to education settings (e.g., Giles et al., 2016). Staff in higher education recognize the importance of test anxiety and express a desire to help but do not recognize the main symptoms and are also unsure how to help (Connon et al., 2016). Other common concerns raised by educators about new interventions are the ease of implementation and the availability of resources, including staffing (Giles et al., 2016; McGeechan et al., 2018). What then, are the implications for integrating new interventions for test anxiety into a higher education setting? There are two major approaches to be considered. The first approach consists of preventive oriented interventions (which could include redesign of the curriculum and assessment), whilst the second is to offer support or interventions to test anxious students. These two approaches are not mutually exclusive. In either case, an important first step will be to explore the views of the educators,

and in particular the students, to ensure engagement with any prevention or intervention programme.

Test anxiety is experienced to at least some degree by all students (Knappe et al., 2011). Therefore, a universally available intervention such as an online-delivered programme based upon the S-REF model would have greatest impact and not place additional burden on teaching staff. Prior to launching such as intervention, it would be important to pilot it and evaluate its feasibility, acceptance, and efficacy. Students should still be able to accessible traditional services for their test anxiety (e.g., face-to-face interventions delivered via wellbeing services). Here, a test-anxiety tailored intervention based on MCT could be offered. The availability of an intervention manual would help ensure consistent delivery. MCT training programmes should also be available to wellbeing services. How might interventions be integrated into undergraduate curricula?

Prevention programmes may consist of many facets. For example, ‘learning and assessment’ sessions discussing effective and evidence-based learning strategies, as well as mental health coping strategies, could be integrated into the curriculum. Students frequently report tests as being an unfair assessment of their abilities (Bonaccio & Reeve, 2010). To address these concerns, two steps could be taken. First, educators may consider multi-modal assessments, such as a mixture of project work, essay assignments, and examinations, to gain more complete picture of a student’s ability. Second, it may be fruitful to explain the examination process in detail to students to demonstrate the care taken in the selection of questions, their wording and psychometric qualities (e.g., for multiple choice questions), and how examination

pass marks are determined. Raising awareness of test anxiety and the avenues available to get support and treatment (e.g., wellbeing services) could also be highlighted in these ‘learning and assessment’ sessions mentioned above and as part of the mental wellbeing weeks that now are commonplace in UK universities.

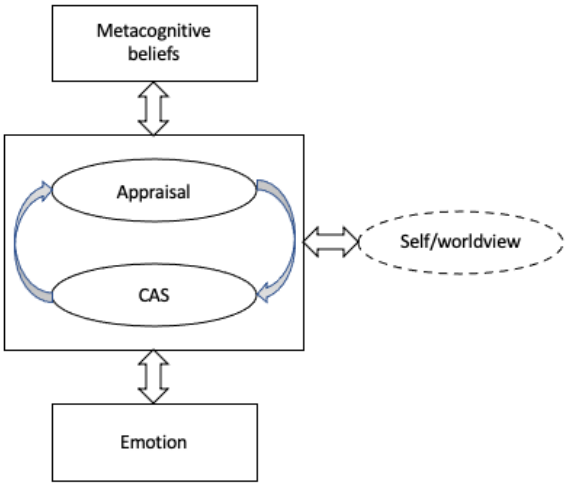
Monitoring and identifying students who may need support for test anxiety can also be achieved via traditional channels, such as through the academic advisors and pastoral care that universities typically offer. Following mock or formative examinations, it would be opportune for academic tutors to discuss learning and test anxiety with their students (even well performing students) and signpost support (e.g., online interventions) or wellbeing services if needed.

If traditional face-to-face interventions are offered, they will need to be delivered in a timely manner to be helpful to students. For example, in a typical UK university, formative results are often only available from February, with summative examinations running around June, and Easter in-between. This only leaves 8-10 weeks available for picking up of referrals, recruitment into an intervention programme, and the delivery of the intervention itself. University wellbeing services report being stretched and frequently have waiting lists (Hubble & Bolton, 2020; Rückert, 2015). Therefore, it is important to identify test anxious students and their needs early if they are to access support prior to undertaking important examinations.

7.7 Concluding remarks

The S-REF model proposes that maladaptive metacognitive beliefs critical to the initiation and maintenance of emotional disorder. Evidence presented here supports the role of metacognitive beliefs in test anxiety. Should future research provide further support for the S-REF model then it may be fruitful to apply an intervention based on metacognitive therapy (MCT), which is derived from the S-REF model, to test anxiety. MCT focuses on removing the CAS, modifying maladaptive metacognitive beliefs, and training students in new ways of thinking and relating to their thoughts. A generic S-REF model case formulation has been proposed (see Figure 7.2).

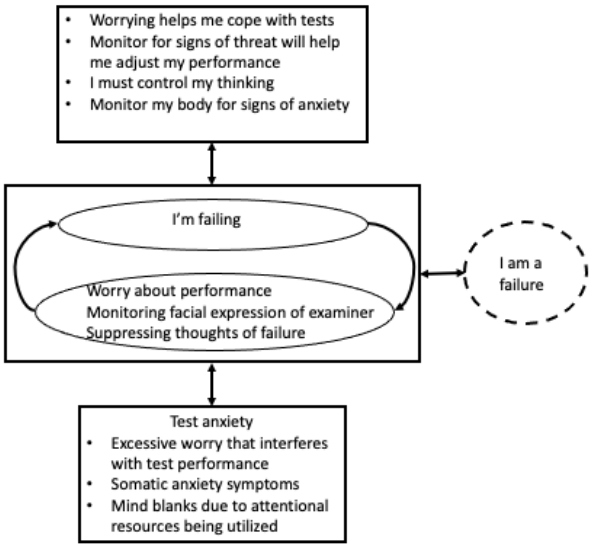
Figure 7.2. Universal S-REF case formulation.



Here, metacognitive beliefs primarily refer to both positive beliefs about the use of worry and engaging in threat monitoring and other coping strategies, and negative beliefs about the uncontrollability and harm of perseverative thinking. The CAS consists of worry and rumination, threat monitoring, and

unhelpful coping strategies. Emotions consists of affect (e.g., anxiety) and its associated physical (e.g., sweating) and behavioural (e.g., restlessness) symptoms. The self / worldview refers to the content of beliefs or thoughts in the ordinary cognitive domain (e.g., the world is a dangerous place) that shapes the content of worry and rumination but is not involved in initiation, continuation, or termination of those processes. An example of how this formulation may apply to test anxiety in an OSCE is provided in Figure 7.3. In this example, test anxiety is amplified and maintained by the CAS through worry about the consequences of failure, attentional focus on anxiety symptoms and scanning memory for gaps, as well as maladaptive coping strategies such as thought suppression and distraction.

Figure 7.3. Example S-REF case formulation for a test-anxious medical student.



Thus far a specific version of the S-REF model has been adapted to each specific disorder (e.g., Generalized Anxiety Disorder). It will be necessary to establish if specific modifications to the generic S-REF model are required for

test anxiety, or indeed whether other beliefs such as IU should also be challenged and modified in an intervention. Interventions that reduce test anxiety will help students achieve their academic potential. Test anxiety is associated with other mental health problems such as depression (Akinsola & Nwajei, 2013) so there may be wider benefits of interventions like MCT. Successful reduction of test anxiety will improve the examination performance and the mental wellbeing of students.

References

Abado, E., Aue, T., & Okon-Singer, H. (2020). The Missing Pieces of the Puzzle: A Review on the Interactive Nature of A-Priori Expectancies and Attention Bias toward Threat. *Brain Sciences, 10*(10), 745. Retrieved from https://www.mdpi.com/2076-3425/10/10/745

Abdi, H. M., Bageri, S., Shoghi, S., Goodarzi, S., & Hosseinzadeh, A. (2012). The role of metacognitive and self-efficacy beliefs in students' test anxiety and academic achievement. *Australian Journal of Basic and Applied Sciences, 6*(12), 418-422. Retrieved from https://www.scopus.com/inward/record.uri?eid=2-s2.0-84872689961&partnerID=40&md5=c1d7610e287d9d8607399f484bc60821

Addis, M. E., Cardemil, E. V., Duncan, B. L., & Miller, S. D. (2006). Does Manualization Improve Therapy Outcomes? In *Evidence-based practices in mental health: Debate and dialogue on the fundamental questions.* (pp. 131-160). Washington, DC, US: American Psychological Association.

Akinsola, E. F., & Nwajei, A. D. (2013). Test Anxiety, Depression and Academic Performance: Assessment and Management Using Relaxation and Cognitive Restructuring Techniques. *Psychology, 4*, 18-24. doi:10.4236/psych.2013.46A1003

Allen, G. J. (1971). Effectiveness of study counseling and desensitization in alleviating test anxiety in college students. *Journal of Abnormal Psychology, 77*(3), 282-289. doi:10.1037/h0031021

Allen, G. J. (1973). Treatment of test anxiety by group-administered and self-administered relaxation and study counseling. *Behavior Therapy, 4*(3), 349-360. doi:https://doi.org/10.1016/S0005-7894(73)80115-7

Allen, G. J., & Desaulniers, G. J. (1974). Effectiveness of study counseling and desensitization in alleviating test anxiety: A two-year follow-up. *Journal of Abnormal Psychology, 83*(2), 186-191. doi:10.1037/h0036482

Alpert, R., & Haber, R. N. (1960). ANXIETY IN ACADEMIC-ACHIEVEMENT SITUATIONS. *Journal of Abnormal and Social Psychology, 61*(2), 207-215. doi:10.1037/h0045464

Altmaier, E. M., & Woodward, M. (1981). Group vicarious desensitization of test anxiety *Journal of Counseling Psychology, 28*(5), 467-469. doi:10.1037/0022-0167.28.5.467

Anastasi, A. (1981). Coaching, test sophistication, and developed abilities. *American Psychologist, 36*(10), 1086–1093.

Anderson, S. B., & Sauser Jr., W. I. (1995). Measurement of test anxiety: An overview. In C. D. S. P. R. V. (Eds.) (Ed.), *Test anxiety: Theory, assessment, and treatment (pp. 15–34)*. Washington, DC: Taylor & Francis.

Angelidis, A., Solis, E., Lautenbach, F., van der Does, W., & Putman, P. (2019). I'm going to fail! Acute cognitive performance anxiety increases threat-interference and impairs WM performance. *Plos One, 14*(2), e0210824. doi:10.1371/journal.pone.0210824

Anton, W. D. (1976). Evaluation of outcome variables in systematic desensitization of test anxiety *Behaviour Research and Therapy, 14*(3), 217-224. doi:10.1016/0005-7967(76)90014-0

Ardito, R., & Rabellino, D. (2011). Therapeutic Alliance and Outcome of Psychotherapy: Historical Excursus, Measurements, and Prospects for Research. *Frontiers in Psychology, 2*(270). doi:10.3389/fpsyg.2011.00270

Association, A. P. (2013). *Diagnostic and statistical manual of mental disorders (5th ed.).*

Association, A. P. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington DC: Author.

Baddeley, A. (2012). Working Memory: Theories, Models, and Controversies. In S. T. Fiske, D. L. Schacter, & S. E. Taylor (Eds.), *Annual Review of Psychology, Vol 63* (Vol. 63, pp. 1-29).

Bailey, R., & Wells, A. (2015). Metacognitive beliefs moderate the relationship between catastrophic misinterpretation and health anxiety. *J Anxiety Disord, 34*, 8-14. doi:10.1016/j.janxdis.2015.05.005

Bailey, R., & Wells, A. (2016). Is metacognition a causal moderator of the relationship between catastrophic misinterpretation and health anxiety? A prospective study. *Behaviour Research and Therapy, 78*, 43-50. doi:https://doi.org/10.1016/j.brat.2016.01.002

Bandura, A. (1965). Influence of models' reinforcement contingencies on the acquisition of imitative responses. *Journal of Personality and Social Psychology, 1*(6), 589-595. doi:10.1037/h0022070

Bar-Haim, Y., Lamy, D., Pergamin, L., Bakermans-Kranenburg, M. J., & van, I. M. H. (2007). Threat-related attentional bias in anxious and nonanxious individuals: a meta-analytic study. *Psychol Bull, 133*(1), 1-24. doi:10.1037/0033-2909.133.1.1

Barabasz, A. F., & Barabasz, M. (1981). Effects of rational-emotive therapy on psychophysiological and reported measures of test anxiety arousal

Journal of Clinical Psychology, 37(3), 511-514. doi:10.1002/1097-4679(198107)37:3<511::aid-jclp2270370311>3.0.co;2-3

Baranoff, J., & Oei, T. P. S. (2015). The cognitive content-specificity hypothesis: Contributions to diagnosis and assessment. In *Assessment in cognitive therapy.* (pp. 172-196). New York, NY, US: The Guilford Press.

Bauman, W., & Melnyk, W. T. (1994). A controlled comparison of eye-movements and finger tapping in the treatment of test anxiety *Journal of Behavior Therapy and Experimental Psychiatry, 25*(1), 29-33. doi:10.1016/0005-7916(94)90060-4

Beck, A. T. (1967). *Depression: Causes and treatment*. Philadelphia: University of Pennsylvania Press.

Beck, A., Emery, G., & Greenberg, R. (1985). *Anxiety Disorders and Phobias: A Cognitive Perspective*. New York: Basic Books.

Beck, A. T. (1976). *Cognitive therapy and the emotional disorders*. Oxford, England: International Universities Press.

Bedell, J. R. (1976). Systematic desensitization, relaxation training and suggestion in the treatment anxiety *Behaviour Research and Therapy, 14*(4), 309-311. doi:10.1016/0005-7967(76)90008-5

Benjamin, M., McKeachie, W. J., Lin, Y. G., & Holinger, D. P. (1981). TEST ANXIETY - DEFICITS IN INFORMATION-PROCESSING. *Journal of Educational Psychology, 73*(6), 816-824. Retrieved from <Go to ISI>://WOS:A1981MU74800006

Biasi, V., Cerutti, R., Mallia, L., Menozzi, F., Patrizi, N., & Violani, C. (2017). (Mal)Adaptive Psychological Functioning of Students Utilizing

University Counseling Services. *Frontiers in Psychology, 8*. doi:10.3389/fpsyg.2017.00403

Bieling, P. J., Antony, M. M., & Swinson, R. P. (1998). The State-Trait Anxiety Inventory, Trait version: structure and content re-examined. *Behaviour Research and Therapy, 36*(7-8), 777-788. doi:10.1016/s0005-7967(98)00023-0

Birrell, J., Meares, K., Wilkinson, A., & Freeston, M. (2011). Toward a definition of intolerance of uncertainty: A review of factor analytical studies of the Intolerance of Uncertainty Scale. *Clinical Psychology Review, 31*(7), 1198-1208. doi:10.1016/j.cpr.2011.07.009

Bishop, S. (2009). Bishop SJ. Trait anxiety and impoverished prefrontal control of attention. Nat Neurosci. 2009;12(1):92-8. *Nature neuroscience, 12*, 92-98. doi:10.1038/nn.2242

Boelen, P. A., & Reijntjes, A. (2009). Intolerance of uncertainty and social anxiety. *Journal of Anxiety Disorders, 23*(1), 130-135. doi:10.1016/j.janxdis.2008.04.007

Bonaccio, S., & Reeve, C. L. (2010). The nature and relative importance of students' perceptions of the sources of test anxiety. *Learning and Individual Differences, 20*(6), 617-625. doi:10.1016/j.lindif.2010.09.007

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. (2009). *Introduction to meta-analysis*. Chichester, U.K.: John Wiley & Sons.

Borkovec, T. D. (1994). The nature, functions and origins of worry. In D. G. C. & T. F. (Eds.), *Worrying: Perspectives on Theory, As- sessment and Treatment* (pp. 5-33). Chichester, UK: Wiley.

Borkovec, T. D., Alcaine, O. M., & Behar, E. (2004). Avoidance Theory of Worry and Generalized Anxiety Disorder. In *Generalized anxiety disorder: Advances in research and practice.* (pp. 77-108). New York, NY, US: The Guilford Press.

Borkovec, T. D., Hazlett-Stevens, H., & Diaz, M. L. (1999). The role of positive beliefs about worry in generalized anxiety disorder and its treatment. *Clinical Psychology & Psychotherapy, 6*(2), 126-138. doi:https://doi.org/10.1002/(SICI)1099-0879(199905)6:2<126::AID-CPP193>3.0.CO;2-M

Borkovec, T. D., & Hu, S. (1990). THE EFFECT OF WORRY ON CARDIOVASCULAR-RESPONSE TO PHOBIC IMAGERY. *Behaviour Research and Therapy, 28*(1), 69-73. doi:10.1016/0005-7967(90)90056-o

Borkovec, T. D., & Inz, J. (1990). THE NATURE OF WORRY IN GENERALIZED ANXIETY DISORDER - A PREDOMINANCE OF THOUGHT ACTIVITY. *Behaviour Research and Therapy, 28*(2), 153-158. doi:10.1016/0005-7967(90)90027-g

Borkovec, T. D., Lyonfields, J. D., Wiser, S. L., & Deihl, L. (1993). THE ROLE OF WORRISOME THINKING IN THE SUPPRESSION OF CARDIOVASCULAR-RESPONSE TO PHOBIC IMAGERY. *Behaviour Research and Therapy, 31*(3), 321-324. doi:10.1016/0005-7967(93)90031-o

Borkovec, T. D., Robinson, E., Pruzinsky, T., & Depree, J. A. (1983). PRELIMINARY EXPLORATION OF WORRY - SOME CHARACTERISTICS AND PROCESSES. *Behaviour Research and Therapy, 21*(1), 9-16. doi:10.1016/0005-7967(83)90121-3

Boswell, J. F., Thompson-Hollands, J., Farchione, T. J., & Barlow, D. H. (2013). Intolerance of uncertainty: a common factor in the treatment of emotional disorders. *J Clin Psychol, 69*(6), 630-645. doi:10.1002/jclp.21965

Bottesi, G., Ghisi, M., Carraro, E., Barclay, N., Payne, R., & Freeston, M. H. (2016). Revising the Intolerance of Uncertainty Model of Generalized Anxiety Disorder: Evidence from UK and Italian Undergraduate Samples. *Frontiers in Psychology, 7*. doi:10.3389/fpsyg.2016.01723

Bottesi, G., Noventa, S., Freeston, M. H., & Ghisi, M. (2019). Seeking certainty about Intolerance of Uncertainty: Addressing old and new issues through the Intolerance of Uncertainty Scale-Revised. *Plos One, 14*(2). doi:10.1371/journal.pone.0211929

Brand, H. S., & Schoonheim-Klein, M. (2009). Is the OSCE more stressful? Examination anxiety and its consequences in different assessment methods in dental education. *European Journal of Dental Education, 13*(3), 147-153. doi:10.1111/j.1600-0579.2008.00554.x

Brewer, C. (1972). Beneficial effect of beta-adrenergic blockade on "exam nerves" *The Lancet, 300*(7774), 435. doi:https://doi.org/10.1016/S0140-6736(72)91840-5

Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York: Guildford.

Browne, M. W., & Cudeck, R. (1992). Alternative Ways of Assessing Model Fit. *Sociological Methods & Research, 21*(2), 230-258. doi:10.1177/0049124192021002005

Bruch, M. A. (1981). Relationship of test-taking strategies to test anxiety and performance: Toward a task analysis of examination behavior. *Cognitive Therapy and Research, 5*(1), 41-56. doi:10.1007/BF01172325

Bruch, M. A., Juster, H. R., & Kaflowitz, N. G. (1983). Relationships of cognitive components of test anxiety to test performance: Implications for assessment and treatment. *Journal of Counseling Psychology, 30*(4), 527-536. doi:10.1037/0022-0167.30.4.527

Buhr, K., & Dugas, M. J. (2002). The intolerance of uncertainty scale: psychometric properties of the English version. *Behaviour Research and Therapy, 40*(8), 931-945. doi:10.1016/s0005-7967(01)00092-4

Buhr, K., & Dugas, M. J. (2006). Investigating the construct validity of intolerance of uncertainty and its unique relationship with worry. *Journal of Anxiety Disorders, 20*(2), 222-236. doi:10.1016/j.janxdis.2004.12.004

Buhr, K., & Dugas, M. J. (2009). The role of fear of anxiety and intolerance of uncertainty in worry: An experimental manipulation. *Behaviour Research and Therapy, 47*(3), 215-223. doi:10.1016/j.brat.2008.12.004

Buhr, K., & Dugas, M. J. (2012). Fear of Emotions, Experiential Avoidance, and Intolerance of Uncertainty in Worry and Generalized Anxiety Disorder. *International Journal of Cognitive Therapy, 5*(1), 1-17. Retrieved from <Go to ISI>://WOS:000301313000001

Buratti, S., & Allwood, C. M. (2013). The effects of advice and "try more" instructions on improving realism of confidence. *Acta Psychologica, 144*(1), 136-144. doi:10.1016/j.actpsy.2013.05.008

Buratti, S., Allwood, C. M., & Kleitman, S. (2013). First- and second-order metacognitive judgments of semantic memory reports: The influence of

personality traits and cognitive styles. *Metacognition and Learning, 8*(1), 79-102. doi:10.1007/s11409-013-9096-5

Burcas, S., & Cretu, R. Z. (2021). Multidimensional Perfectionism and Test Anxiety: a Meta-analytic Review of Two Decades of Research. *Educational Psychology Review, 33*(1), 249-273. doi:10.1007/s10648-020-09531-3

C.S., C., & M.F., S. (1991). Self-Regulation and the Self. In S. J. & G. G.R. (Eds.), *The Self: Interdisciplinary Approaches*. New York: Springer.

Calvo, M. G., Eysenck, M. W., & Castillo, M. D. (1997). Interpretation bias in test anxiety: The time course of predictive inferences. *Cognition & Emotion, 11*(1), 43-63. doi:10.1080/026999397380023

Carleton, R. N. (2012). The intolerance of uncertainty construct in the context of anxiety disorders: theoretical and practical perspectives. *Expert Review of Neurotherapeutics, 12*(8), 937-947. doi:10.1586/ern.12.82

Carleton, R. N. (2016). Into the unknown: A review and synthesis of contemporary models involving uncertainty. *Journal of Anxiety Disorders, 39*, 30-43. doi:10.1016/j.janxdis.2016.02.007

Carleton, R. N., Collimore, K. C., & Asmundson, G. J. G. (2010). "It's not just the judgements-It's that I don't know": Intolerance of uncertainty as a predictor of social anxiety. *Journal of Anxiety Disorders, 24*(2), 189-195. doi:10.1016/j.janxdis.2009.10.007

Carleton, R. N., Gosselin, P., & Asmundson, G. J. G. (2010). The Intolerance of Uncertainty Index: Replication and Extension With an English Sample. *Psychological Assessment, 22*(2), 396-406. doi:10.1037/a0019230

Carleton, R. N., Mulvogue, M. K., Thibodeau, M. A., McCabe, R. E., Antony, M. M., & Asmundson, G. J. G. (2012). Increasingly certain about uncertainty: Intolerance of uncertainty across anxiety and depression. *Journal of Anxiety Disorders, 26*(3), 468-479. doi:10.1016/j.janxdis.2012.01.011

Carleton, R. N., Norton, P. J., & Asmundson, G. J. G. (2007). Fearing the unknown: A short version of the Intolerance of Uncertainty Scale. *Journal of Anxiety Disorders, 21*(1), 105-117. doi:10.1016/j.janxdis.2006.03.014

Carleton, R. N., Weeks, J. W., Howell, A. N., Asmundson, G. J. G., Antony, M. M., & McCabe, R. E. (2012). Assessing the latent structure of the intolerance of uncertainty construct: An initial taxometric analysis. *Journal of Anxiety Disorders, 26*(1), 150-157. doi:10.1016/j.janxdis.2011.10.006

Cartwright-Hatton, S., & Wells, A. (1997). Beliefs about worry and intrusions: the Meta-Cognitions Questionnaire and its correlates. *J Anxiety Disord, 11*(3), 279-296. doi:10.1016/s0887-6185(97)00011-x

Carver, C. S., Peterson, L. M., Follansbee, D. J., & Scheier, M. F. (1983). EFFECTS OF SELF-DIRECTED ATTENTION ON PERFORMANCE AND PERSISTENCE AMONG PERSONS HIGH AND LOW IN TEST ANXIETY. *Cognitive Therapy and Research, 7*(4), 333-354. doi:10.1007/bf01177556

Cassady, J. C. (2004a). The impact of cognitive test anxiety on text comprehension and recall in the absence of external evaluative pressure. *Applied Cognitive Psychology, 18*(3), 311-325. doi:10.1002/acp.968

Cassady, J. C. (2004b). The influence of cognitive test anxiety across the learning-testing cycle. *Learning and Instruction, 14*(6), 569-592. doi:10.1016/j.learninstruc.2004.09.002

Cassady, J. C. (2010). Test anxiety: Contemporary theories and implications for learning. In J. C. Cassady (Ed.), *Anxiety in schools: The causes, consequences, and solutions for academic anxieties* (pp. 7-26). New York, NY: Peter Lang.

Cassady, J. C., & Finch, W. H. (2015). Using factor mixture modeling to identify dimensions of cognitive test anxiety. *Learning and Individual Differences, 41*, 14-20. doi:10.1016/j.lindif.2015.06.002

Cassady, J. C., & Finch, W. H. (2020). Revealing Nuanced Relationships Among Cognitive Test Anxiety, Motivation, and Self-Regulation Through Curvilinear Analyses. *Frontiers in Psychology, 11*, 1141-1141. doi:10.3389/fpsyg.2020.01141

Cassady, J. C., & Johnson, R. E. (2002). Cognitive test anxiety and academic performance. *Contemporary Educational Psychology, 27*(2), 270-295. doi:10.1006/ceps.2001.1094

Cassady, J. C., Mohammed, A., & Mathieu, L. (2004). Cross-cultural differences in test perceptions - Women in Kuwait and the United States. *Journal of Cross-Cultural Psychology, 35*(6), 713-718. doi:10.1177/0022022104270113

Cerasoli, C. P., Nicklin, J. M., & Ford, M. T. (2014). Intrinsic Motivation and Extrinsic Incentives Jointly Predict Performance: A 40-Year Meta-Analysis. *Psychological Bulletin, 140*(4), 980-1008. doi:10.1037/a0035661

Chan, A.-W., Tetzlaff, J. M., Gøtzsche, P. C., Altman, D. G., Mann, H., Berlin, J. A., . . . Moher, D. (2013). SPIRIT 2013 explanation and elaboration: guidance for protocols of clinical trials. *BMJ : British Medical Journal, 346*, e7586. doi:10.1136/bmj.e7586

Chang-Liang, R., & Denney, D. R. (1976). Applied relaxation as training in self-control *Journal of Counseling Psychology, 23*(3), 183-189. doi:10.1037//0022-0167.23.3.183

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. doi:10.1037/0022-0663.97.2.268

Chen, H. (2012). Impact of Parent's Socioeconomic Status on Perceived Parental Pressure and Test Anxiety among Chinese High School Students. *International Journal of Psychological Studies, 4*(234-245).

Chen, R., Zhou, R. L., & Ismbe. (2008). *An Experimental Research on Attentional Disengagement in Test-anxious Students*.

Cho, Y., Jahng, S., & Chai, S. (2012). The Factor Structure and Concurrent Validity of the Korean Version of the Metacognitions Questionnaire 30 (K-MCQ-30). *Journal of Clinical Psychology, 68*(3). doi:10.1002/jclp.20867

Chukwuorji, J. C., & Nwonyi, S. K. (2015). Test anxiety: contributions of gender, age, parent's occupation and self-esteem among secondary school students in Nigeria. *Journal of Psychology in Africa, 25*(1), 60-64. doi:10.1080/14330237.2015.1007600

Cisler, J. M., & Koster, E. H. (2010). Mechanisms of attentional biases towards threat in anxiety disorders: An integrative review. *Clin Psychol Rev, 30*(2), 203-216. doi:10.1016/j.cpr.2009.11.003

Clark, D. M., & Wells, A. (1995). A cognitive model of social phobia. In *Social phobia: Diagnosis, assessment, and treatment.* (pp. 69-93). New York, NY, US: The Guilford Press.

Clawson, T. W., Firment, C. K., & Trower, T. L. (1981). TEST ANXIETY - ANOTHER ORIGIN FOR RACIAL BIAS IN STANDARDIZED TESTING. *Measurement and Evaluation in Guidance, 13*(4), 210-215. doi:10.1080/00256307.1981.12022238

Cohen, J. (1988). *Statistical power analysis for the behavioural sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*(1), 155-159. doi:10.1037/0033-2909.112.1.155

Collaboration, T. C. (2011). Cochrane Handbook for Systematic Reviews of Interventions: Online version (5.1.0, March 2011). In. Available from www.cochrane-handbook.org.

Collaboration, T. C. (2014). Review Manager (RevMan) [Computer program]. Version 5.3.5. Copenhagen, Denmark The Nordic Cochrane Centre.

Connon, H. A., Rash, J. A., Allen Gerwing, A. M., Bramble, B., Landine, J., & Gerwing, T. G. (2016). Post-Secondary Educators' Perceptions of Students' Test Anxiety. *The Canadian Journal for the Scholarship of Teaching and Learning, 7*. doi:https://doi.org/10.5206/cjsotl-rcacea.2016.1.9

Cook, D. J., Mulrow, C. D., & Haynes, R. B. (1997). Systematic reviews: synthesis of best evidence for clinical decisions. *Ann Intern Med, 126*(5), 376-380. doi:10.7326/0003-4819-126-5-199703010-00006

Cook, S. A., Salmon, P., Dunn, G., & Fisher, P. (2014). Measuring metacognition in cancer: Validation of the metacognitions questionnaire 30 (MCQ-30). *Plos One, 9*(9). doi:10.1371/journal.pone.0107302

Cook, S. A., Salmon, P., Dunn, G., Holcombe, C., Cornford, P., & Fisher, P. (2015). The association of metacognitive beliefs with emotional distress after diagnosis of cancer. *Health Psychol, 34*(3), 207-215. doi:10.1037/hea0000096

Covington, M. V., & Omelich, C. L. (1988). Achievement dynamics: The interaction of motives, cognitions, and emotions over time. *Anxiety Research, 1*(3), 165-183. doi:10.1080/08917778808248717

Cornacchio, D., Sanchez, A. L., Coxe, S., Roy, A., Pincus, D. B., Read, K. L., . . . Comer, J. S. (2018). Factor structure of the intolerance of uncertainty scale for children. *Journal of Anxiety Disorders, 53*, 100-107. doi:10.1016/j.janxdis.2017.07.003

Counsell, A., Furtado, M., Iorio, C., Anand, L., Canzonieri, A., Fine, A., . . . Katzman, M. A. (2017). Intolerance of uncertainty, social anxiety, and generalized anxiety: Differences by diagnosis and symptoms. *Psychiatry Res, 252*, 63-69. doi:10.1016/j.psychres.2017.02.046

Counts, D. K., Hollandsworth, J. G., & Alcorn, J. D. (1978). USE OF ELECTROMYOGRAPHIC BIOFEEDBACK AND CUE-CONTROLLED RELAXATION IN TREATMENT OF TEST

ANXIETY. *Journal of Consulting and Clinical Psychology, 46*(5), 990-996. doi:10.1037/0022-006x.46.5.990

Covington, M. V., & Omelich, C. L. (1988). Achievement dynamics: The interaction of motives, cognitions, and emotions over time. *Anxiety Research, 1*(3), 165-183. doi:10.1080/08917778808248717

Craig, P., Dieppe, P., Macintyre, S., Michie, S., Nazareth, I., & Petticrew, M. (2008). Developing and evaluating complex interventions: the new Medical Research Council guidance. *BMJ, 337*, a1655. doi:10.1136/bmj.a1655

Crouse, R. H., Deffenbacher, J. L., & Frost, G. A. (1985). DESENSITIZATION FOR STUDENTS WITH DIFFERENT SOURCES AND EXPERIENCES OF TEST ANXIETY. *Journal of College Student Development, 26*(4), 315-318. Retrieved from <Go to ISI>://WOS:A1985AKU4300005

Crowley, C., Crowley, D., & Clodfelter, C. (1986). Effects of a self-coping cognitive treatment for test anxiety *Journal of Counseling Psychology, 33*(1), 84-86. doi:10.1037//0022-0167.33.1.84

Culler, R. E., & Holahan, C. J. (1980). TEST ANXIETY AND ACADEMIC-PERFORMANCE - EFFECTS OF STUDY-RELATED BEHAVIORS. *Journal of Educational Psychology, 72*(1), 16-20. doi:10.1037/0022-0663.72.1.16

D'Alelio, W. A., & Murray, E. J. (1981). Cognitive therapy for test anxiety *Cognitive Therapy and Research, 5*(3), 299-307. doi:10.1007/bf01193413

Darvishzadeh, P., Aguilar, M., & Moradi, A.-R. (2012). A comparative study of working memory executive functions processing efficiency considering

high and low levels of anxiety. *Procedia - Social and Behavioral Sciences, 32*, 40–44. doi:10.1016/j.sbspro.2012.01.007

David, D., & Szentagotai, A. (2006). Cognitions in cognitive-behavioral psychotherapies; toward an integrative model. *Clinical Psychology Review, 26*(3), 284-298. doi:10.1016/j.cpr.2005.09.003

Dawley, H. H. (1973). GROUP IMPLOSIVE THERAPY IN TREATMENT OF TEST ANXIETY - BRIEF REPORT. *Behavior Therapy, 4*(2), 261-263. doi:10.1016/s0005-7894(73)80036-x

Dawley, H. H., & Wenrich, W. W. (1973a). MASSED GROUP DESENSITIZATION IN REDUCTION OF TEST-ANXIETY. *Psychological Reports, 33*(2), 359-363. Retrieved from <Go to ISI>://WOS:A1973R168400005

Dawley, H. H., & Wenrich, W. W. (1973b). Treatment of test anxiety by group implosive therapy *Psychological Reports, 33*(2), 383-388. Retrieved from <Go to ISI>://WOS:A1973R168400011

Decker, T. W., & Russell, R. K. (1981). Comparison of cue-controlled relaxation and cognitive restructuring versus study skills counseling in treatment of test-anxious college underachievers *Psychological Reports, 49*(2), 459-469. Retrieved from <Go to ISI>://WOS:A1981MP64700022

Decker, M. L., Turk, C. L., Hess, B., & Murray, C. E. (2008). Emotion regulation among individuals classified with and without generalized anxiety disorder. *J Anxiety Disord, 22*(3), 485-494. doi:10.1016/j.janxdis.2007.04.002

Deffenbacher, J. L. (1974). HIERARCHIES FOR DESENSITIZATION OF TEST AND SPEECH ANXIETIES. *Journal of College Student*

Development, *15*(6), 452-454. Retrieved from <Go to ISI>://WOS:A1974U663200005

Deffenbacher, J. L. (1976). RELAXATION INVIVO IN TREATMENT OF TEST ANXIETY. *Journal of Behavior Therapy and Experimental Psychiatry*, *7*(3), 289-292. doi:10.1016/0005-7916(76)90017-3

Deffenbacher, J. L. (1978). WORRY, EMOTIONALITY, AND TASK-GENERATED INTERFERENCE IN TEST ANXIETY - EMPIRICAL-TEST OF ATTENTIONAL THEORY. *Journal of Educational Psychology*, *70*(2), 248-254. doi:10.1037/0022-0663.70.2.248

Deffenbacher, J. L. (1980). Worry and emotionality in test anxiety. In I. G. Sarason (Ed.), *Test anxiety: Theory, research and application*. Hillsdale, New Jersey: Erlbaum.

Deffenbacher, J. L. (1986). COGNITIVE AND PHYSIOLOGICAL COMPONENTS OF TEST ANXIETY IN REAL-LIFE EXAMS. *Cognitive Therapy and Research*, *10*(6), 635-644. doi:10.1007/bf01173751

Deffenbacher, J. L., & Deitz, S. R. (1978). EFFECTS OF TEST ANXIETY ON PERFORMANCE, WORRY, AND EMOTIONALITY IN NATURALLY OCCURRING EXAMS. *Psychology in the Schools*, *15*(3), 446-450. doi:10.1002/1520-6807(197807)15:3<446::Aid-pits2310150326>3.0.Co;2-b

Deffenbacher, J. L., & Hazaleus, S. L. (1985). COGNITIVE, EMOTIONAL, AND PHYSIOLOGICAL COMPONENTS OF TEST ANXIETY. *Cognitive Therapy and Research*, *9*(2), 169-180. doi:10.1007/bf01204848

Deffenbacher, J. L., Mathis, H., & Michaels, A. C. (1979). Two self-control procedures in the reduction of targeted and nontargeted anxities *Journal of Counseling Psychology, 26*(2), 120-127. doi:10.1037//0022-0167.26.2.120

Deffenbacher, J. L., & Michaels, A. C. (1980). Two self-control procedures in the reduction of targeted and nontargeted anxities - a year later *Journal of Counseling Psychology, 27*(1), 9-15. Retrieved from <Go to ISI>://WOS:A1980JC59000002

Deffenbacher, J. L., & Michaels, A. C. (1981a). A 12-month follow-up of homogeneous and heterogeneous anxiety management training *Journal of Counseling Psychology, 28*(5), 463-466. doi:10.1037//0022-0167.28.5.463

Deffenbacher, J. L., & Michaels, A. C. (1981b). Anxiety management training and self-control desensitization - 15-months later *Journal of Counseling Psychology, 28*(5), 459-462. doi:10.1037/0022-0167.28.5.459

Deffenbacher, J. L., Michaels, A. C., Daley, P. C., & Michaels, T. (1980). A comparison of homogenuous and heterogeneous anxiety management training *Journal of Counseling Psychology, 27*(6), 630-634. Retrieved from <Go to ISI>://WOS:A1980KQ05300015

Deffenbacher, J. L., Michaels, A. C., Michaels, T., & Daley, P. C. (1980). Comparison of anxiety management training and self-control desensitization *Journal of Counseling Psychology, 27*(3), 232-239. doi:10.1037/0022-0167.27.3.232

Deffenbacher, J. L., & Parks, D. H. (1979). A comparison of traditional and self-control desensitization *Journal of Counseling Psychology, 26*(2), 93-97. doi:10.1037/0022-0167.26.2.93

Deffenbacher, J. L., & Rivera, N. (1976). BEHAVIORAL SELF-CONTROL TREATMENT OF TEST ANXIETY IN MINORITY POPULATIONS - SOME CASES AND ISSUES. *Psychological Reports, 39*(3), 1188-1190. doi:10.2466/pr0.1976.39.3f.1188

Deffenbacher, J. L., & Shelton, J. L. (1978). COMPARISON OF ANXIETY MANAGEMENT-TRAINING AND DESENSITIZATION IN REDUCING TEST AND OTHER ANXIETIES. *Journal of Counseling Psychology, 25*(4), 277-282. Retrieved from <Go to ISI>://WOS:A1978FH07500005

Deffenbacher, J. L., & Snyder, A. L. (1976). RELAXATION AS SELF-CONTROL IN TREATMENT OF TEST AND OTHER ANXIETIES. *Psychological Reports, 39*(2), 379-385. Retrieved from <Go to ISI>://WOS:A1976CJ66000009

DeLucia-Waack, J. L., Kalodner, C. R., & Riva, M. T. (2013). *Handbook of Group Counseling and Psychotherapy (2nd edition)*. London: Sage.

Dendato, K. M., & Diener, D. (1986). Effectiveness of cognitive/relaxation therapy and study-skills training in reducing self-reported anxiety and improving the academic performance of test-anxious students. *Journal of Counseling Psychology, 33*(2), 131-135. doi:10.1037/0022-0167.33.2.131

Denney, D. R., & Rupert, P. A. (1977). Desensitization and self-control in the treatment of test anxiety *Journal of Counseling Psychology, 24*(4), 272-280. doi:10.1037//0022-0167.24.4.272

Depreeuw, E., & DeNeve, H. (1992). Test anxiety can harm your health: Some conclusions based on a student typology. In D. G. Forgays, T. Sosnowski, & K. Wrzesniewski (Eds.), *Anxiety: Recent developments in cognitive, psycho physiological, and health research* (pp. 211-228). Washington Hemisphere.

Derakshan, N., & Eysenck, M. W. (2009). Anxiety, Processing Efficiency, and Cognitive Performance New Developments from Attentional Control Theory. *European Psychologist, 14*(2), 168-176. doi:10.1027/1016-9040.14.2.168

Derakshan, N., Smyth, S., & Eysenck, M. W. (2009). Effects of state anxiety on performance using a task-switching paradigm: an investigation of attentional control theory. *Psychon Bull Rev, 16*(6), 1112-1117. doi:10.3758/pbr.16.6.1112

DerSimonian, R., & Laird, N. (1986). Meta-analysis in clinical trials. *Controlled Clinical Trials, 7*(3), 177-188.

DiStefano, C., Liu, J., Jiang, N., & Shi, D. (2018). Examination of the Weighted Root Mean Square Residual: Evidence for Trustworthiness? *Structural Equation Modeling: A Multidisciplinary Journal, 25*(3), 453-466. doi:10.1080/10705511.2017.1390394

Doctor, D. M., & Altman, F. (1969). WORRY AND EMOTIONALITY AS COMPONENTS OF TEST ANXIETY - REPLICATION AND FURTHER DATA. *Psychological Reports, 24*(2), 563-+. Retrieved from <Go to ISI>://WOS:A1969D552000048

Doctor, R. M., & Altman, F. (1969). Worry and emotionality as components of test anxiety: Replication and further data. *Psychological Reports, 24*(2), 563-568. doi:10.2466/pr0.1969.24.2.563

Dugas, M. J., Brillon, P., Savard, P., Turcotte, J., Gaudet, A., Ladouceur, R., . . . Gervais, N. J. (2010). A randomized clinical trial of cognitive-behavioral therapy and applied relaxation for adults with generalized anxiety disorder. *Behav Ther, 41*(1), 46-58. doi:10.1016/j.beth.2008.12.004

Dugas, M. J., Freeston, M. H., & Ladouceur, R. (1997). Intolerance of uncertainty and problem orientation in worry. *Cognitive Therapy and Research, 21*(6), 593-606. doi:10.1023/a:1021890322153

Dugas, M. J., Gagnon, F., Ladouceur, R., & Freeston, M. H. (1998). Generalized anxiety disorder: a preliminary test of a conceptual model. *Behaviour Research and Therapy, 36*(2), 215-226. doi:10.1016/s0005-7967(97)00070-3

Dugas, M. J., Gosselin, P., & Ladouceur, R. (2001). Intolerance of uncertainty and worry: Investigating specificity in a nonclinical sample. *Cognitive Therapy and Research, 25*(5), 551-558. doi:10.1023/a:1005553414688

Dugas, M. J., Hedayati, M., Karavidas, A., Buhr, K., Francis, K., & Phillips, N. A. (2005). Intolerance of Uncertainty and Information Processing: Evidence of Biased Recall and Interpretations. *Cognitive Therapy and Research, 29*(1), 57-70. doi:10.1007/s10608-005-1648-9

Dugas, M. J., & Ladouceur, R. (2000). Treatment of GAD - Targeting intolerance of uncertainty in two types of worry. *Behavior Modification, 24*(5), 635-657. doi:10.1177/0145445500245002

Dugas, M. J., Ladouceur, R., Leger, E., Freeston, M. H., Langlois, F., Provencher, M. D., & Boisvert, J. M. (2003). Group cognitive-behavioral therapy for generalized anxiety disorder: Treatment outcome and long-term follow-up. *Journal of Consulting and Clinical Psychology, 71*(4), 821-825. doi:10.1037/0022-006x.71.4.821

Dugas, M. J., Marchand, A., & Ladouceur, R. (2005). Further validation of a cognitive-behavioral model of generalized anxiety disorder: diagnostic and symptom specificity. *Journal of Anxiety Disorders, 19*(3), 329-343. doi:10.1016/j.janxdis.2004.02.002

Dugas, M. J., Savard, P., Gaudet, A., Turcotte, J., Laugesen, N., Robichaud, M., . . . Koerner, N. (2007). Can the components of a cognitive model predict the severity of generalized anxiety disorder? *Behav Ther, 38*(2), 169-178. doi:10.1016/j.beth.2006.07.002

Dutke, S., & Stober, J. (2001). Test anxiety, working memory, and cognitive performance: Supportive effects of sequential demands. *Cognition & Emotion, 15*(3), 381-389. doi:10.1080/0269993004200231

Efthimiou, O., Debray, T. P., van Valkenhoef, G., Trelle, S., Panayidou, K., Moons, K. G., . . . Salanti, G. (2016). GetReal in network meta-analysis: a review of the methodology. *Res Synth Methods, 7*(3), 236-263. doi:10.1002/jrsm.1195

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal, 315*(7109), 629-634. Retrieved from <Go to ISI>://WOS:A1997XW69400016

Elliot, A. J., & McGregor, H. A. (1999). Test anxiety and the hierarchical model of approach and avoidance achievement motivation. *Journal of Personality and Social Psychology, 76*(4), 628-644. doi:10.1037//0022-3514.76.4.628

Elliot, A. J., & McGregor, H. A. (2001). A 2 x 2 achievement goal framework. *Journal of Personality and Social Psychology, 80*(3), 501-519. doi:10.1037//0022-3514.80.3.501

Ellis, A. (1957). Rational Psychotherapy and Individual Psychology. *Journal of Individual Psychology, 13*, 38-44.

Ellis, A. (1962). *Reason and Emotion in Psychotherapy*. New York, NY: Stuart.

Endler, N. S., & Parker, J. D. A. (1990a). MULTIDIMENSIONAL ASSESSMENT OF COPING - A CRITICAL-EVALUATION. *Journal of Personality and Social Psychology, 58*(5), 844-854. doi:10.1037/0022-3514.58.5.844

Endler, N. S., & Parker, J. D. A. (1990b). STATE AND TRAIT ANXIETY, DEPRESSION AND COPING STYLES. *Australian Journal of Psychology, 42*(2), 207-220. doi:10.1080/00049539008260119

Ergene, T. (2003). Effective interventions on test anxiety reduction - A meta-analysis. *School Psychology International, 24*(3), 313-328. doi:10.1177/01430343030243004

Everson, H. T., Millsap, R. E., & Rodriguez, C. M. (1991). Isolating gender differences in test anxiety: A confirmatory factor-analysis of the test anxiety inventory *Educational and Psychological Measurement, 51*(1), 243-251. doi:10.1177/0013164491511024

Eysenck, M. W. (1985). ANXIETY AND COGNITIVE-TASK PERFORMANCE. *Personality and Individual Differences, 6*(5), 579-586. doi:10.1016/0191-8869(85)90007-8

Eysenck, M. W. (1997). *Anxiety and cognition: A unified theory*. Hove, England: Psychology Press/Erlbaum (UK) Taylor & Francis.

Eysenck, M. W., & Byrne, A. (1994). IMPLICIT MEMORY BIAS, EXPLICIT MEMORY BIAS, AND ANXIETY. *Cognition & Emotion, 8*(5), 415-431. doi:10.1080/02699939408408950

Eysenck, M. W., & Calvo, M. G. (1992). ANXIETY AND PERFORMANCE - THE PROCESSING EFFICIENCY THEORY. *Cognition & Emotion, 6*(6), 409-434. doi:10.1080/02699939208409696

Eysenck, M. W., & Derakshan, N. (2011). New perspectives in attentional control theory. *Personality and Individual Differences, 50*(7), 955-960. doi:10.1016/j.paid.2010.08.019

Eysenck, M. W., Derakshan, N., Santos, R., & Calvo, M. G. (2007). Anxiety and cognitive performance: Attentional control theory. *Emotion, 7*(2), 336-353. doi:10.1037/1528-3542.7.2.336

Fergus, T. A., Bardeen, J. R., & Wu, K. D. (2013). Intolerance of uncertainty and uncertainty-related attentional biases: Evidence of facilitated engagement or disengagement difficulty? *Cognitive Therapy and Research, 37*(4), 735-741. doi:10.1007/s10608-012-9509-9

Fergus, T. A., & Carleton, R. N. (2016). Intolerance of uncertainty and attentional networks: Unique associations with alerting. *Journal of Anxiety Disorders, 41*, 59-64. doi:10.1016/j.janxdis.2016.03.010

Fergus, T. A., & Wheless, N. (2018). Examining incremental explanatory power in accounting for worry severity: negative metacognitive beliefs uniquely predict worry severity following a worry episode. *Anxiety Stress and Coping, 31*(5), 514-525. doi:10.1080/10615806.2018.1479828

Fergus, T. A., & Wu, K. D. (2013). The Intolerance of Uncertainty Scale: Measurement Invariance, Population Heterogeneity, and Its Relation With Worry Among Self-Identifying White and Black Respondents. *Assessment, 20*(5), 555-564. doi:10.1177/1073191112460272

Fernie, B. A., Kopar, U. Y., Fisher, P. L., & Spada, M. M. (2018). Further development and testing of the metacognitive model of procrastination: Self-reported academic performance. *Journal of Affective Disorders, 240*, 1-5. doi:10.1016/j.jad.2018.07.018

Field, A. P. (2009). *Discovering statistics using SPSS*. London, UK: Sage.

Finger, R., & Galassi, J. P. (1977a). Effects of modifying cognitive versus emotionality responses in the treatment of test anxiety *Journal of Consulting and Clinical Psychology, 45*(2), 280-287. doi:10.1037//0022-006x.45.2.280

Finger, R., & Galassi, J. P. (1977b). EFFECTS OF MODIFYING COGNITIVE VERSUS EMOTIONALITY RESPONSES IN TREATMENT OF TEST ANXIETY. *Journal of Consulting and Clinical Psychology, 45*(2), 280-287. doi:10.1037/0022-006x.45.2.280

Fisher, P. L. (2006). The effectiveness of psychological treatments for generalised anxiety disorder. In A. W. G. C. L. Davey (Ed.), *Worry and its psychological disorders: Assessment and treatment* (pp. 200-220). Chichester: John Wiley & Sons Ltd.

Fisher, P. L., Cook, S. A., & Noble, A. (2016). Clinical utility of the Metacognitions Questionnaire 30 in people with epilepsy. *Epilepsy and Behavior, 57*, 185-191. doi:10.1016/j.yebeh.2016.02.004

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American Psychologist, 34*(10), 906-911. doi:10.1037/0003-066X.34.10.906

Fleur, D. S., Bredeweg, B., & van den Bos, W. (2021). Metacognition: ideas and insights from neuro- and educational sciences. *NPJ Sci Learn, 6*(1), 13. doi:10.1038/s41539-021-00089-5

Freemantle, N., & Geddes, J. (1998). Understanding and interpreting systematic reviews and meta-analyses. Part 2: meta-analyses. *Evidence Based Mental Health, 1*(4), 102-104. doi:10.1136/ebmh.1.4.102

Freeston, M. H., Rheaume, J., Letarte, H., Dugas, M. J., & Ladouceur, R. (1994). Why do people worry. *Personality and Individual Differences, 17*(6), 791-802. doi:10.1016/0191-8869(94)90048-5

Freud, S. (1899/2011). *The Interpretation of Dreams*. London: Basic Books.

Friedman, I. A., & Bendas-Jacob, O. (1997). Measuring perceived test anxiety in adolescents: A self-report scale. *Educational and Psychological Measurement, 57*(6), 1035-1046. doi:10.1177/0013164497057006012

Galassi, J. P., Frierson, H. T., & Sharer, R. (1981). Behavior of high, moderate, and low test anxious students during an actual test situation. *Journal of Consulting and Clinical Psychology, 49*(1), 51-62. doi:10.1037/0022-006X.49.1.51

Galassi, J. P., Frierson, H. T., & Sharer, R. (1981). CONCURRENT VERSUS RETROSPECTIVE ASSESSMENT IN TEST ANXIETY RESEARCH.

Journal of Consulting and Clinical Psychology, 49(4), 614-615. doi:10.1037/0022-006x.49.4.614

Galassi, J. P., Frierson, H. T., & Siegel, R. G. (1984). COGNITIONS, TEST ANXIETY, AND TEST-PERFORMANCE - A CLOSER LOOK. *Journal of Consulting and Clinical Psychology, 52*(2), 319-320. doi:10.1037/0022-006x.52.2.319

Gallagher, J. W., & Arkowitz, H. (1978). Weak effects of covert modeling treatment of test anxiety *Journal of Behavior Therapy and Experimental Psychiatry, 9*(1), 23-26. doi:10.1016/0005-7916(78)90083-6

Garlington, W. K., & Cotler, S. B. (1968). Systematic desensitization of test anxiety. *Behaviour Research and Therapy, 6*(3), 247-256.

Gaudry, E., Vagg, P., & Spielberger, C. D. (1975). VALIDATION OF STATE-TRAIT DISTINCTION IN ANXIETY RESEARCH. *Multivariate Behavioral Research, 10*(3), 331-341. doi:10.1207/s15327906mbr1003_6

Geen, R. G. (1976). TEST ANXIETY, OBSERVATION, AND RANGE OF CUE UTILIZATION. *British Journal of Social and Clinical Psychology, 15*(SEP), 253-259. doi:10.1111/j.2044-8260.1976.tb00032.x

Geen, R. G. (1985). TEST ANXIETY AND VISUAL VIGILANCE. *Journal of Personality and Social Psychology, 49*(4), 963-970. doi:10.1037/0022-3514.49.4.963

Geen, R. G. (1987). TEST ANXIETY AND BEHAVIORAL AVOIDANCE. *Journal of Research in Personality, 21*(4), 481-488. doi:10.1016/0092-6566(87)90034-1

Gerwing, T. G., Rash, J. A., Gerwing, A. M. A., Bramble, B., & Landine, J. (2015). Perceptions and incidence of test anxiety. *The Canadian Journal*

for the Scholarship of Teaching and Learning, *6*(3). doi:http://dx.doi.org/10.5206/cjsotl-rcacea.2015.3.3

Giles, E. L., Coulton, S., Deluca, P., Drummond, C., Howel, D., Kaner, E., . . . Newbury-Birch, D. (2016). Multicentre individual randomised controlled trial of screening and brief alcohol intervention to prevent risky drinking in young people aged 14-15 in a high school setting (SIPS JR-HIGH): study protocol. *BMJ Open*, *6*(12), e012474. doi:10.1136/bmjopen-2016-012474

Glass, G. V., McGaw, B., & L., S. M. (1981). *Meta-analysis in social research*. Newbury Park: Sage Publications.

Goldacre, B. (2015). How to Get All Trials Reported: Audit, Better Data, and Individual Accountability. *PLOS Medicine*, *12*(4), e1001821. doi:10.1371/journal.pmed.1001821

Gormley, G. (2011). Summative OSCEs in undergraduate medical education. *The Ulster medical journal*, *80*(3), 127-132. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/23526843

https://www.ncbi.nlm.nih.gov/pmc/PMC3605523/

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3605523/pdf/umj0080-0127.pdf

Gosselin, P., Ladouceur, R., Evers, A., Laverdiere, A., Routhier, S., & Tremblay-Picard, M. (2008). Evaluation of intolerance of uncertainty: Development and validation of a new self-report measure. *Journal of Anxiety Disorders*, *22*(8), 1427-1439. doi:10.1016/j.janxdis.2008.02.005

Green, K. E. (1981). TEST ANXIETY LEVEL AND TEST FORMAT PREFERENCE. *Psychological Reports*, *48*(2), 537-538. doi:10.2466/pr0.1981.48.2.537

Gros, D. F., Antony, M. M., Simms, L. J., & McCabe, R. E. (2007). Psychometric properties of the State-Trait Inventory for Cognitive and Somatic Anxiety (STICSA): Comparison to the State-Trait Anxiety Inventory (STAI). *Psychological Assessment, 19*(4), 369-381. doi:10.1037/1040-3590.19.4.369

Grøtte, T., Solem, S., Myers, S. G., Hjemdal, O., Vogel, P. A., Güzey, I. C., . . . Fisher, P. (2016). Metacognitions in obsessive-compulsive disorder: A psychometric study of the Metacognitions Questionnaire-30. *Journal of Obsessive-Compulsive and Related Disorders, 11*, 82-90. doi:10.1016/j.jocrd.2016.09.002

Guida, F. V., & Ludlow, L. H. (1989). A CROSS-CULTURAL-STUDY OF TEST ANXIETY. *Journal of Cross-Cultural Psychology, 20*(2), 178-190. doi:10.1177/0022022189202004

Guidry, L. S., & Randolph, D. L. (1974). Covert reinforcement in the treatment of test anxiety. *Journal of Counseling Psychology, 21*(4), 260-264. doi:10.1037/h0036724

Guraya, S. Y., Guraya, S. S., Habib, F., AlQuiliti, K. W., & Khoshhal, K. I. (2018). Medical students' perception of test anxiety triggered by different assessment modalities. *Medical Teacher, 40*, S49-S55. doi:10.1080/0142159x.2018.1465178

Guyatt, G. H., Oxman, A. D., Kunz, R., Woodcock, J., Brozek, J., Helfand, M., . . . Grp, G. W. (2011). GRADE guidelines: 7. Rating the quality of evidence-inconsistency. *Journal of Clinical Epidemiology, 64*(12), 1294-1302. doi:10.1016/j.jclinepi.2011.03.017

Guyatt, G. H., Oxman, A. D., Vist, G. E., Kunz, R., Falck-Ytter, Y., Alonso-Coello, P., . . . Grp, G. W. (2008). GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *British Medical Journal, 336*(7650), 924-926. doi:10.1136/bmj.39489.470347.AD

Hahm, N., Augustin, S., Bade, C., Ammer-Wies, A., & Bahramsoltani, M. (2016). Test Anxiety: Evaluation of a Low-Threshold Seminar-Based Intervention for Veterinary Students. *Journal of Veterinary Medical Education, 43*(1), 47-57. doi:10.3138/jvme.0215-029R1

Hahn, H., Kropp, P., Kirschstein, T., Rucker, G., & Muller-Hilke, B. (2017). Test anxiety in medical school is unrelated to academic performance but correlates with an effort/reward imbalance. *Plos One, 12*(2). doi:10.1371/journal.pone.0171220

Hahne, R., Lohmann, R., & Krzyszycha, K. (1999). *Studium und psychische Probleme: Sonderauswertung zur 15 Sozialerhebung des Deutschen Studentenwerks*. Bonn: Deutsches Studentenwerk.

Hale, W., Richmond, M., Bennett, J., Berzins, T., Fields, A., Weber, D., . . . Osman, A. (2016). Resolving Uncertainty About the Intolerance of Uncertainty Scale-12: Application of Modern Psychometric Strategies. *Journal of Personality Assessment, 98*(2), 200-208. doi:10.1080/00223891.2015.1070355

Hall, T. S. (2005). *Is test anxiety a form of specific social phobia?* (PhD Thesis). University of Maryland,

Harden, R. M. (1988). What is an OSCE? *Medical Teacher, 10*(1), 19-22. doi:10.3109/01421598809019321

Hebert, E. A., & Dugas, M. J. (2019). Behavioral Experiments for Intolerance of Uncertainty: Challenging the Unknown in the Treatment of Generalized Anxiety Disorder. *Cognitive and Behavioral Practice, 26*(2), 421-436. Retrieved from <Go to ISI>://WOS:000466826300015

Hedges, L. V. (1981). Distribution Theory for Glass's Estimator of Effect Size and Related Estimators. *Journal of Educational Statistics, 6*(2), 107-128. doi:10.2307/1164588

Hedges, L. V., & Olkin, I. (1980). Vote-counting methods in research synthesis. *Psychological Bulletin, 88*(2), 359-369. doi:10.1037/0033-2909.88.2.359

Helton, W. S., & Näswall, K. (2015). Short Stress State Questionnaire: Factor structure and state change assessment. *European Journal of Psychological Assessment, 31*(1), 20-30. doi:10.1027/1015-5759/a000200

Hembree, R. (1988). Correlates, causes, effects, and treatment of test anxiety *Review of Educational Research, 58*(1), 47-77. doi:10.3102/00346543058001047

Herbert, J. D., & Forman, E. M. (2011). The evolution of cognitive behavior therapy: The rise of psychological acceptance and mindfulness. In J. D. Herbert & E. M. Forman (Eds.), *Acceptance and mindfulness in cognitive behavior therapy: Understanding and applying the new therapies.* (pp. 3-25). Hoboken, NJ, US: John Wiley & Sons Inc.

Herzer, F., Wendt, J., & Hamm, A. O. (2014). Discriminating clinical from nonclinical manifestations of test anxiety: A validation study. *Behavior Therapy, 45*(2), 222-231. doi:10.1016/j.beth.2013.11.001

Herzer, F., Wendt, J., & Hamm, A. O. (2015). Discriminant validity of constructs derived from the self-regulative model for evaluation anxiety for

predicting clinical manifestations of test anxiety. *Behaviour Research and Therapy, 73*, 52-57. doi:10.1016/j.brat.2015.07.012

Higgins, J. P. T., Altman, D. G., Gøtzsche, P. C., Jüni, P., Moher, D., Oxman, A. D., . . . Sterne, J. A. C. (2011). The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ (Online), 343*(7829). doi:10.1136/bmj.d5928

Higgins, J. P. T., & Green, S. (Eds.). (2011). *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]*. The Cochrane Collaboration: Available from www.cochrane-handbook.org.

Higgins, J. P. T., & Thompson, S. G. (2004). Controlling the risk of spurious findings from meta-regression. *Statistics in Medicine, 23*(11), 1663-1682. doi:10.1002/sim.1752

Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *British Medical Journal, 327*(7414), 557-560. doi:10.1136/bmj.327.7414.557

Hill, K. T., & Wigfield, A. (1984). Test anxiety: A major educational problem and what can be done about it *Elementary School Journal, 85*(1), 105-126. doi:10.1086/461395

Hodapp, V. (1995). The TAI-G: A multidimensional approach to the assessment of test anxiety. In C. Schwarzer & M. Zeidner (Eds.), *Stress, anxiety, and coping in academic settings* (pp. 95-130). Tübingen, Germany: Francke.

Hodapp, V., & Benson, J. (1997). The multidimensionality of test anxiety: A test of different models. *Anxiety, Stress and Coping, 10*(3), 219-244. doi:10.1080/10615809708249302

Hodges, B. D. (2006). The objective structured clinical examination: Three decades of development. *Journal of Veterinary Medical Education, 33*(4), 571-577. doi:10.3138/jvme.33.4.571

Hoffart, A., Johnson, S. U., Nordahl, H. M., & Wells, A. (2018). Mechanisms of change in metacognitive and cognitive behavioral therapy for treatment-resistant anxiety: The role of metacognitive beliefs and coping strategies. *Journal of Experimental Psychopathology, 9*(3). doi:10.1177/2043808718787414

Hofmann, S. G., & Hay, A. C. (2018). Rethinking avoidance: Toward a balanced approach to avoidance in treating anxiety disorders. *J Anxiety Disord, 55*, 14-21. doi:10.1016/j.janxdis.2018.03.004

Holahan, C. J., Richardson, F. C., Puckett, S. P., & Bell, K. F. (1979). Evaluation of two test-anxiety reduction treatments in a secondary prevention program *American Journal of Community Psychology, 7*(6), 679-687. doi:10.1007/bf00891970

Hollandsworth, J. G., Glazeski, R. C., Kirkland, K., Jones, G. E., & Vannorman, L. R. (1979). ANALYSIS OF THE NATURE AND EFFECTS OF TEST ANXIETY - COGNITIVE, BEHAVIORAL, AND PHYSIOLOGICAL COMPONENTS. *Cognitive Therapy and Research, 3*(2), 165-180. doi:10.1007/bf01172603

Holroyd, K. A. (1976). COGNITION AND DESENSITIZATION IN GROUP TREATMENT OF TEST ANXIETY. *Journal of Consulting and Clinical Psychology, 44*(6), 991-1001. doi:10.1037/0022-006x.44.6.991

Holroyd, K. A. (1978). EFFECTIVENESS OF AN ATTRIBUTION THERAPY MANIPULATION WITH TEST ANXIETY. *Behavior Therapy, 9*(4), 526-534. doi:10.1016/s0005-7894(78)80125-7

Holroyd, K. A., Westbrook, T., Wolf, M., & Badhorn, E. (1978). PERFORMANCE, COGNITION, AND PHYSIOLOGICAL RESPONDING IN TEST ANXIETY. *Journal of Abnormal Psychology, 87*(4), 442-451. doi:10.1037/0021-843x.87.4.442

Hong, E. (1998). Differential stability of individual differences in state and trait test anxiety. *Learning and Individual Differences, 10*(1), 51-69. doi:10.1016/s1041-6080(99)80142-3

Hong, E. (1999). Test anxiety, perceived test difficulty, and test performance: Temporal patterns of their effects. *Learning and Individual Differences, 11*(4), 431-447. doi:10.1016/s1041-6080(99)80012-0

Hong, E., & Karstensson, L. (2002). Antecedents of state test anxiety. *Contemporary Educational Psychology, 27*(2), 348-367. doi:10.1006/ceps.2001.1095

Hong, R. Y., & Lee, S. S. M. (2015). Further clarifying prospective and inhibitory intolerance of uncertainty: Factorial and construct validity of test scores from the intolerance of uncertainty scale. *Psychological Assessment, 27*(2), 605-620. doi:10.1037/pas0000074

Horne, A. M., & Matson, J. L. (1977). Comparison of modeling, desensitization, flooding, study skills, and control groups for reducing test anxiety *Behavior Therapy, 8*(1), 1-8. doi:10.1016/s0005-7894(77)80114-7

Howard, E. (2020). *A review of the literature concerning anxiety for educational assessments*. https://www.gov.uk/government/publications/a-review-of-the-literature-on-anxiety-for-educational-assessments: Ofqual

Hu, L. T., & Bentler, P. M. (1999). Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria Versus New Alternatives. *Structural Equation Modeling-a Multidisciplinary Journal, 6*(1), 1-55. doi:10.1080/10705519909540118

Huang, H.-T. D., & Hung, S.-T. A. (2013). Comparing the Effects of Test Anxiety on Independent and Integrated Speaking Test Performance. *TESOL Quarterly, 47*(2), 244-269. doi:https://doi.org/10.1002/tesq.69

Hubble, S., & Bolton, P. (2020). *Support for students with mental health issues in England*. London, UK: House of Commons Library.

Huberty, T. J., & Dick, A. C. (2006). Performance and Test Anxiety. In *Children's needs III: Development, prevention, and intervention.* (pp. 281-291). Washington, DC, US: National Association of School Psychologists.

Hunsley, J. (1985). Test anxiety, academic performance, and cognitive appraisals. *Journal of Educational Psychology, 77*(6), 678-682. doi:10.1037/0022-0663.77.6.678

Huntley, C. D., & Fisher, P. L. (2016). Examining the role of positive and negative metacognitive beliefs in depression. *Scand J Psychol, 57*(5), 446-452. doi:10.1111/sjop.12306

Huntley, C. D., Young, B., Jha, V., & Fisher, P. L. (2016). The efficacy of interventions for test anxiety in university students: A protocol for a

systematic review and meta-analysis. *International Journal of Educational Research, 77*, 92-98. doi:10.1016/j.ijer.2016.03.001

Huntley, C. D., Young, B., Smith, C. T., & Fisher, P. L. (2020). Uncertainty and test anxiety: Psychometric properties of the Intolerance of Uncertainty Scale-12 (IUS-12) among university students. *International Journal of Educational Research, 104*. doi:10.1016/j.ijer.2020.101672

Huntley, C. D., Young, B., Smith, C. T., Jha, V., & Fisher, P. L. (2020). Assessing metacognitive beliefs in test anxiety: Psychometric properties of the metacognitions questionnaire, 30 (MCQ-30) among university students. *Current Psychology*. doi:10.1007/s12144-020-00662-y

Huntley, C. D., Young, B., Temple, J., Longworth, M., Smith, C. T., Jha, V., & Fisher, P. L. (2019). The efficacy of interventions for test-anxious university students: A meta-analysis of randomized controlled trials. *Journal of Anxiety Disorders, 63*, 36-50. doi:10.1016/j.janxdis.2019.01.007

Huntley, C. D., Young, B., Tudur Smith, C., & Fisher, P. L. (2019). Psychometric properties of the Intolerance of Uncertainty - 12 (IUS-12) and Metacognitions Questionnaire - 30 (MCQ-30) in test anxiety. *Manuscript submitted for publication*.

Huntley, C. D., Young, B., Tudur Smith, C., & Fisher, P. L. (submitted). Testing times: The contribution of intolerance of uncertainty and metacognitive beliefs to test anxiety in university students. *Manuscript submitted for publication.*

IBM. (2017). SPSS Statistics for Windows, Version 25.0. Armonk, NY: IBM Corp.

Ihli, K. L., & Garlington, W. K. (1969). A COMPARISON OF GROUP VS INDIVIDUAL DESENSITIZATION OF TEST ANXIETY. *Behaviour Research and Therapy, 7*(2), 207-+. doi:10.1016/0005-7967(69)90035-7

Ikeda, M., Iwanaga, M., & Seiwa, H. (1996). Test anxiety and working memory system. *Perceptual and Motor Skills, 82*(3, Pt 2), 1223-1231. doi:10.2466/pms.1996.82.3c.1223

Jacobson, E. (1938). *Progressive relaxation* (2nd ed.). Oxford, UK: University of Chicago Press.

Jacobson, N. S., Follette, W. C., & Revenstorf, D. (1984). Psychotherapy outcome research - methods for reporting variability and evaluating clinical significance *Behavior Therapy, 15*(4), 336-352. doi:10.1016/s0005-7894(84)80002-7

Jacobson, N. S., & Truax, P. (1991). Clinical significance - a statistical approach to defining meaningful change in psychotherapy research *Journal of Consulting and Clinical Psychology, 59*(1), 12-19. doi:10.1037//0022-006x.59.1.12

Jacoby, R. J., Fabricant, L. E., Leonard, R. C., Riemann, B. C., & Abramowitz, J. S. (2013). Just to be certain: Confirming the factor structure of the Intolerance of Uncertainty Scale in patients with obsessive-compulsive disorder. *Journal of Anxiety Disorders, 27*(5), 535-542. doi:10.1016/j.janxdis.2013.07.008

Janke, S., Rudert, S. C., Marksteiner, T., & Dickhauser, O. (2017). Knowing One's Place: Parental Educational Background Influences Social Identification with Academia, Test Anxiety, and Satisfaction with

Studying at University. *Frontiers in Psychology, 8*. doi:10.3389/fpsyg.2017.01326

Kaplan, R. M., McCordick, S. M., & Twitchell, M. (1979). IS IT THE COGNITIVE OR THE BEHAVIORAL COMPONENT WHICH MAKES COGNITIVE-BEHAVIOR MODIFICATION EFFECTIVE IN TEST ANXIETY. *Journal of Counseling Psychology, 26*(5), 371-377. doi:10.1037/0022-0167.26.5.371

Karpicke, J. D., & Roediger, H. L. (2008). The critical importance of retrieval for learning. *Science, 319*(5865), 966-968. doi:10.1126/science.1152408

Kaur, A., Butow, P., & Thewes, B. (2011). Do Metacognitions Predict Attentional Bias in Health Anxiety? *Cognitive Therapy and Research, 35*(6), 575-580. doi:10.1007/s10608-011-9387-6

Kavakci, O., Semiz, M., Kartal, A., Dikici, A., & Kugu, N. (2014). Test anxiety prevalance and related variables in the students who are going to take the university entrance examination. *Dusunen Adam: The Journal of Psychiatry and Neurological Sciences, 27*, 301-307. doi:10.5350/DAJPN2014270403

Kazdin, A. E. (2015). Evidence-based psychotherapies II: changes in models of treatment and treatment delivery. *South African Journal of Psychology, 45*(1), 3-21. doi:10.1177/0081246314538733

Keech, J. J., Hagger, M. S., O'Callaghan, F. V., & Hamilton, K. (2018). The Influence of University Students' Stress Mindsets on Health and Performance Outcomes. *Annals of Behavioral Medicine, 52*(12), 1046-1059. doi:10.1093/abm/kay008

Kiekens, G., Claes, L., Demyttenaere, K., Auerbach, R. P., Green, J. G., Kessler, R. C., . . . Bruffaerts, R. (2016). Lifetime and 12-Month Nonsuicidal Self-Injury and Academic Performance in College Freshmen. *Suicide and Life-Threatening Behavior, 46*(5), 563-576. doi:10.1111/sltb.12237

Kim, K. J. (2016). Factors associated with medical student test anxiety in objective structured clinical examinations: a preliminary study. *International Journal of Medical Education, 7*, 424-427. doi:10.5116/ijme.5845.caec

Kim, S. H., & Rocklin, T. (1994). THE TEMPORAL PATTERNS OF WORRY AND EMOTIONALITY AND THEIR DIFFERENTIAL-EFFECTS ON TEST-PERFORMANCE. *Anxiety Stress and Coping, 7*(2), 117-130. doi:10.1080/10615809408249339

King, N. J., & Ollendick, T. H. (1989). CHILDRENS ANXIETY AND PHOBIC DISORDERS IN SCHOOL SETTINGS - CLASSIFICATION, ASSESSMENT, AND INTERVENTION ISSUES. *Review of Educational Research, 59*(4), 431-470. doi:10.3102/00346543059004431

Kirkland, K., & Hollandsworth, J. G. (1979). TEST ANXIETY, STUDY SKILLS, AND ACADEMIC-PERFORMANCE. *Journal of College Student Development, 20*(5), 431-436. Retrieved from <Go to ISI>://WOS:A1979HS54500011

Kirkland, K., & Hollandsworth, J. G. (1980). Effective test taking: Skills-acquisition versus anxiety-reduction techniques. *Journal of Consulting and Clinical Psychology, 48*(4), 431-439. doi:10.1037/0022-006X.48.4.431

Klinger, E. (1984). A consciousness-sampling analysis of test anxiety and performance. *Journal of Personality and Social Psychology, 47*(6), 1376-1390. doi:10.1037/0022-3514.47.6.1376

Klug, K., Tolgou, T., Schilbach, M., & Rohrmann, S. (2021). Intrusions in test anxiety. *Current Psychology, 40*(5), 2290-2300. doi:10.1007/s12144-019-0167-x

Knappe, S., Beesdo-Baum, K., Fehm, L., Stein, M. B., Lieb, R., & Wittchen, H. U. (2011). Social fear and social phobia types among community youth: Differential clinical features and vulnerability factors. *Journal of Psychiatric Research, 45*(1), 111-120. doi:10.1016/j.jpsychires.2010.05.002

Knight, A., McLellan, L., Jones, M., & Hudson, J. (2014). Pre-treatment Predictors of Outcome in Childhood Anxiety Disorders: A Systematic Review. *Psychopathology Review, a1*(1), 77-129. doi:10.5127/pr.034613

Kostka, M. P., & Galassi, J. P. (1974a). Group systematic desensitization versus covert positive reinforcement in the reduction of test anxiety *Journal of Counseling Psychology, 21*(6), 464-468. doi:10.1037/h0037290

Kostka, M. P., & Galassi, J. P. (1974b). GROUP SYSTEMATIC-DESENSITIZATION VERSUS COVERT POSITIVE REINFORCEMENT IN REDUCTION OF TEST ANXIETY. *Journal of Counseling Psychology, 21*(6), 464-468. doi:10.1037/h0037290

Krauss, A. (2018). Why all randomised controlled trials produce biased results. *Ann Med, 50*(4), 312-322. doi:10.1080/07853890.2018.1453233

Krispenz, A., Gort, C., Schultke, L., & Dickhauser, O. (2019). How to Reduce Test Anxiety and Academic Procrastination Through Inquiry of

Cognitive Appraisals: A Pilot Study Investigating the Role of Academic Self-Efficacy. *Frontiers in Psychology, 10*. doi:10.3389/fpsyg.2019.01917

Krohne, H. W. (1992). Developmental conditions of anxiety and coping: A two-process model of child-rearing effects. In K. A. Hagtvet & B. T. Johnsen (Eds.), *Advances in test anxiety research (Vol. 5, pp. 27–36)*. Lisse, The Netherlands: Swets & Zeitlinger.

Ladouceur, R., Dugas, M. J., Freeston, M. H., Leger, E., Gagnon, F., & Thibodeau, N. (2000). Efficacy of a cognitive-behavioral treatment for generalized anxiety disorder: Evaluation in a controlled clinical trial. *Journal of Consulting and Clinical Psychology, 68*(6), 957-964. doi:10.1037/0022-006x.68.6.957

Ladouceur, R., Dugas, M. J., Freeston, M. H., Rheaume, J., Blais, F., Boisvert, J. M., . . . Thibodeau, N. (1999). Specificity of generalized anxiety disorder symptoms and processes. *Behavior Therapy, 30*(2), 191-207. doi:10.1016/s0005-7894(99)80003-3

Ladouceur, R., Gosselin, P., & Dugas, M. J. (2000). Experimental manipulation of intolerance of uncertainty: a study of a theoretical model of worry. *Behaviour Research and Therapy, 38*(9), 933-941. doi:10.1016/s0005-7967(99)00133-3

Ladouceur, R., Talbot, F., & Dugas, M. J. (1997). Behavioral expressions of intolerance of uncertainty in worry - Experimental findings. *Behavior Modification, 21*(3), 355-371. doi:10.1177/01454455970213006

Lang, K. A., Mueller, J. H., & Nelson, R. E. (1983). TEST ANXIETY AND SELF-SCHEMAS. *Motivation and Emotion, 7*(2), 169-178. doi:10.1007/bf00992900

Lauriola, M., Mosca, O., & Carleton, R. N. (2016). HIERARCHICAL FACTOR STRUCTURE OF THE INTOLERANCE OF UNCERTAINTY SCALE SHORT FORM (IUS-12) IN THE ITALIAN VERSION. *Tpm-Testing Psychometrics Methodology in Applied Psychology, 23*(3), 377-394. doi:10.4473/tpm23.3.8

Lauriola, M., Mosca, O., Trentini, C., Foschi, R., Tambelli, R., & Carleton, R. N. (2018). The Intolerance of Uncertainty Inventory: Validity and Comparison of Scoring Methods to Assess Individuals Screening Positive for Anxiety and Depression. *Frontiers in Psychology, 9*. doi:10.3389/fpsyg.2018.00388

Lay, C. H., Edwards, J. M., Parker, J. D. A., & Endler, N. S. (1989). AN ASSESSMENT OF APPRAISAL, ANXIETY, COPING, AND PROCRASTINATION DURING AN EXAMINATION PERIOD. *European Journal of Personality, 3*(3), 195-208. doi:10.1002/per.2410030305

Lazarus, R. S. (1991). Progress on a cognitive-motivational-relational theory of emotion. *American Psychologist, 46*(8), 819-834. doi:10.1037/0003-066X.46.8.819

Lazarus, R. S., & Folkman, S. (1984). *Stress, appraisal, and coping*. New York: Springer.

LeBeau, R. T., Glenn, D., Liao, B., Wittchen, H. U., Beesdo-Baum, K., Ollendick, T., & Craske, M. G. (2010). SPECIFIC PHOBIA: A REVIEW

OF DSM-IV SPECIFIC PHOBIA AND PRELIMINARY RECOMMENDATIONS FOR DSM-V. *Depression and Anxiety, 27*(2), 148-167. doi:10.1002/da.20655

Lee, J. H. (1999). Test anxiety and working memory. *Journal of Experimental Education, 67*(3), 218-240. doi:10.1080/00220979909598354

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 and Life-Threatening Behavior, 36*(1), 82-96. doi:10.1521/suli.2006.36.1.82

Lent, R. W., Lopez, F. G., Romano, J. L., & Wambach, C. A. (1983). A PROGRAM FOR REDUCING TEST ANXIETY WITH ACADEMICALLY UNDERPREPARED STUDENTS. *Journal of College Student Development, 24*(3), 265-266. Retrieved from <Go to ISI>://WOS:A1983QU71100013

Lent, R. W., & Russell, R. K. (1978). Treatment of test anxiety by cue-controlled desensitization and study-skills training *Journal of Counseling Psychology, 25*(3), 217-224. doi:10.1037//0022-0167.25.3.217

Levine, R. A., & O'brien, R. M. (1980). Treatment of anxiety about college tests with negative practice and systematic desensitization: Some negative findings *Psychological Reports, 46*(3), 823-829. Retrieved from <Go to ISI>://WOS:A1980KE99600030

Levy, H. C., & Radomsky, A. S. (2016). Are all safety behaviours created equal? A comparison of novel and routinely used safety behaviours in obsessive-compulsive disorder. *Cogn Behav Ther, 45*(5), 367-379. doi:10.1080/16506073.2016.1184712

Liao, K. Y., & Wei, M. (2011). Intolerance of uncertainty, depression, and anxiety: the moderating and mediating roles of rumination. *J Clin Psychol*, *67*(12), 1220-1239. doi:10.1002/jclp.20846

Liebert, R. M., & Morris, L. W. (1967). COGNITIVE AND EMOTIONAL COMPONENTS OF TEST ANXIETY - A DISTINCTION AND SOME INITIAL DATA. *Psychological Reports*, *20*(3P1), 975-&. doi:10.2466/pr0.1967.20.3.975

Lin, Y. C., & McKeachie, W. J. (1971). SEX SIMILARITY IN PERSONALITY CORRELATES OF TEST ANXIETY. *Psychological Reports*, *29*(2), 515-+. doi:10.2466/pr0.1971.29.2.515

Linden, A. N., Lau-Barraco, C., & Milletich, R. J. (2014). Protective behavioral strategies, alcohol expectancies, and drinking motives in a model of college student drinking. *Psychology of Addictive Behaviors*, *28*(4), 952-959. doi:10.1037/a0037041

Liu, D. F., & Xu, B. (2017). Test anxiety: perceptions of American community college nursing students. *Empirical Research in Vocational Education and Training*, *9*. doi:10.1186/s40461-017-0048-1

Llera, S. J., & Newman, M. G. (2010). Effects of worry on physiological and subjective reactivity to emotional stimuli in generalized anxiety disorder and nonanxious control participants. *Emotion*, *10*(5), 640-650. doi:10.1037/a0019351

Lomont, J. F., & Sherman, L. J. (1971). Group systematic desensitization and group insight therapies for test anxiety *Behavior Therapy*, *2*(4), 511-518. doi:10.1016/s0005-7894(71)80097-7

Loprinzi, P. D., & Frith, E. (2018). The Role of Sex in Memory Function: Considerations and Recommendations in the Context of Exercise. *Journal of Clinical Medicine, 7*(6). doi:10.3390/jcm7060132

Lotz, C., & Sparfeldt, J. R. (2017). Does test anxiety increase as the exam draws near? - Students' state test anxiety recorded over the course of one semester. *Personality and Individual Differences, 104*, 397-400. doi:10.1016/j.paid.2016.08.032

Lowe, P. A., & Lee, S. W. (2008). Factor Structure of the Test Anxiety Inventory for Children and Adolescents (TAICA) Scores Across Gender Among Students in Elementary and Secondary School Settings. *Journal of Psychoeducational Assessment, 26*(3), 231-246. doi:10.1177/0734282907303773

Lurie, E. S., & Steffen, J. J. (1980). Effective components of covert reinforcement *Journal of Psychology, 106*(2), 241-248. Retrieved from <Go to ISI>://WOS:A1980KS99000008

Mamun, M. A., Misti, J. M., Hosen, I., & Al Mamun, F. (2021). Suicidal behaviors and university entrance test-related factors: A Bangladeshi exploratory study. *Perspect Psychiatr Care*. doi:10.1111/ppc.12783

Mandler, G., & Sarason, S. B. (1952). A study of anxiety and learning. *The Journal of Abnormal and Social Psychology, 47*(2), 166–173.

Marshall, G., & Jones, N. (2003). A pilot study into the anxiety induced by various assessment methods. *Radiography, 9*(3), 185-191. doi:10.1016/S1078-8174(03)00062-2

Mathews, A., May, J., Mogg, K., & Eysenck, M. (1990). ATTENTIONAL BIAS IN ANXIETY - SELECTIVE SEARCH OR DEFECTIVE FILTERING.

Journal of Abnormal Psychology, 99(2), 166-173. doi:10.1037/0021-843x.99.2.166

Matthews, G., Campbell, S. E., Falconer, S., Joyner, L. A., Huggins, J., Gilliland, K., . . . Warm, J. S. (2002). Fundamental Dimensions of Subjective State in Performance Settings: Task Engagement, Distress, and Worry. *Emotion, 2*(4), 315-340. doi:10.1037//1528-3542.2.4.315

Matthews, G., & Harley, T. A. (1996). Connectionist models of emotional distress and attentional bias. *Cognition & Emotion, 10*(6), 561-600. Retrieved from <Go to ISI>://WOS:A1996VV27300001

Matthews, G., Hillyard, E. J., & Campbell, S. E. (1999). Metacognition and maladaptive coping as components of test anxiety. *Clinical Psychology & Psychotherapy, 6*(2), 111-125. Retrieved from <Go to ISI>://WOS:000081114100007

Matthews, G., & Wells, A. (2000). Attention, automaticity, and affective disorder. *Behavior Modification, 24*(1), 69-93. doi:10.1177/0145445500241004

Maxfield, L., & Melnyk, W. T. (2000). Single session treatment of test anxiety with eye movement desensitization and reprocessing (EMDR). *International Journal of Stress Management, 7*(2), 87-101. doi:10.1023/a:1009580101287

May, C., Finch, T., Mair, F., Ballini, L., Dowrick, C., Eccles, M., . . . Heaven, B. (2007). Understanding the implementation of complex interventions in health care: the normalization process model. *BMC Health Services Research, 7*(1), 148. doi:10.1186/1472-6963-7-148

May, C. R., Finch, T. L., Cornford, J., Exley, C., Gately, C., Kirk, S., . . . Mair, F. S. (2011). Integrating telecare for chronic disease management in the community: What needs to be done? *BMC Health Services Research, 11*(1), 131. doi:10.1186/1472-6963-11-131

McCabe, S. E. (2008). Misperceptions of non-medical prescription drug use: a web survey of college students. *Addictive behaviors, 33*(5), 713-724. doi:10.1016/j.addbeh.2007.12.008

McCabe, S. E., Teter, C. J., & Boyd, C. J. (2006). Medical use, illicit use, and diversion of abusable prescription drugs. *Journal of American college health : J of ACH, 54*(5), 269-278. doi:10.3200/JACH.54.5.269-278

McCordick, S. M., Kaplan, R. M., Finn, M. E., & Smith, S. H. (1979). COGNITIVE BEHAVIOR-MODIFICATION AND MODELING FOR TEST ANXIETY. *Journal of Consulting and Clinical Psychology, 47*(2), 419-420. doi:10.1037//0022-006x.47.2.419

McCordick, S. M., Kaplan, R. M., Smith, S., & Finn, M. E. (1981). Variations in cognitive behavior modification for test anxiety *Psychotherapy-Theory Research and Practice, 18*(2), 170-178. doi:10.1037/h0086077

McEvoy, P. M., & Erceg-Hurn, D. M. (2016). The search for universal transdiagnostic and trans-therapy change processes: Evidence for intolerance of uncertainty. *Journal of Anxiety Disorders, 41*, 96-107. doi:10.1016/j.janxdis.2016.02.002

McEvoy, P. M., Hyett, M. P., Shihata, S., Price, J. E., & Strachan, L. (2019). The impact of methodological and measurement factors on transdiagnostic associations with intolerance of uncertainty: A meta-analysis. *Clinical Psychology Review, 73*. doi:10.1016/j.cpr.2019.101778

McEvoy, P. M., & Mahoney, A. E. J. (2011). Achieving certainty about the structure of intolerance of uncertainty in a treatment-seeking sample with anxiety and depression. *Journal of Anxiety Disorders, 25*(1), 112-122. doi:10.1016/j.janxdis.2010.08.010

McGeechan, G. J., Giles, E. L., Scott, S., McGovern, R., Boniface, S., Ramsay, A., . . . Team, t. S. J.-H. S. (2018). A qualitative exploration of school-based staff's experiences of delivering an alcohol screening and brief intervention in the high school setting: findings from the SIPS JR-HIGH trial. *Journal of Public Health, 41*(4), 821-829. doi:10.1093/pubmed/fdy184

McGregor, H. A., & Elliot, A. J. (2002). Achievement goals as predictors of achievement-relevant processes prior to task engagement. *Journal of Educational Psychology, 94*(2), 381-395. doi:10.1037//0022-0663.94.2.381

Meichenbaum, D. H. (1972). Cognitive modification of test anxious college students. *Journal of Consulting and Clinical Psychology, 39*(3), 370-380.

Melnick, J., & Russell, R. W. (1976). Hypnosis versus systematic desensitization in treatment of test anxiety. *Journal of Counseling Psychology, 23*(4), 291-295. doi:10.1037//0022-0167.23.4.291

Mennin, D. S., Heimberg, R. G., Turk, C. L., & Fresco, D. M. (2002). Applying an Emotion Regulation Framework to Integrative Approaches to Generalized Anxiety Disorder. *Clinical Psychology: Science and Practice, 9*, 85-90.

Mennin, D. S., Heimberg, R. G., Turk, C. L., & Fresco, D. M. (2005). Preliminary evidence for an emotion dysregulation model of generalized

anxiety disorder. *Behaviour Research and Therapy, 43*(10), 1281-1310. doi:https://doi.org/10.1016/j.brat.2004.08.008

Meyer, T. J., Miller, M. L., Metzger, R. L., & Borkovec, T. D. (1990). Development and validation of the Penn State Worry Questionnaire *Behaviour Research and Therapy, 28*(6), 487-495. doi:10.1016/0005-7967(90)90135-6

Mitchell, K. R. (1971). EFFECTS OF NEUROTICISM ON INTRA-TREATMENT RESPONSIVITY TO GROUP DESENSITIZATION OF TEST ANXIETY. *Behaviour Research and Therapy, 9*(4), 371-&. doi:10.1016/0005-7967(71)90051-9

Mitchell, K. R., Hall, R. F., & Piatkowska, O. E. (1975a). GROUP PROGRAM FOR BRIGHT FAILING UNDERACHIEVERS. *Journal of College Student Development, 16*(4), 306-312. Retrieved from <Go to ISI>://WOS:A1975AJ31600010

Mitchell, K. R., Hall, R. F., & Piatkowska, O. E. (1975b). A group program for treatment of failing college students *Behavior Therapy, 6*(3), 324-336. doi:10.1016/s0005-7894(75)80107-9

Mitchell, K. R., & Ingham, R. J. (1970). The effects of general anxiety on group desensitization of test anxiety *Behaviour Research and Therapy, 8*(1), 69-&. doi:10.1016/0005-7967(70)90037-9

Mitchell, K. R., & Ng, K. T. (1972). EFFECTS OF GROUP COUNSELING AND BEHAVIOR THERAPY ON ACADEMIC ACHIEVEMENT OF TEST-ANXIOUS STUDENTS. *Journal of Counseling Psychology, 19*(6), 491-&. doi:10.1037/h0033584

Mitchell, K. R., & Piatkows.Oe. (1974). EFFECTS OF GROUP TREATMENT FOR COLLEGE UNDERACHIEVERS AND BRIGHT FAILING UNDERACHIEVERS. *Journal of Counseling Psychology, 21*(6), 494-501. doi:10.1037/h0037312

Moher, D., Hopewell, S., Schulz, K. F., Montori, V., Gøtzsche, P. C., Devereaux, P. J., . . . Altman, D. G. (2010). CONSORT 2010 Explanation and Elaboration: updated guidelines for reporting parallel group randomised trials. *BMJ, 340*, c869. doi:10.1136/bmj.c869

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Grp, P. (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLOS Medicine, 6*(7). doi:10.1371/journal.pmed.1000097

Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., . . . Grp, P.-P. (2016). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Revista Espanola de Nutricion Humana y Dietetica, 20*(2), 148-160. doi:10.14306/renhyd.20.2.223

Moneta, G. B., Spada, M. M., & Rost, F. M. (2007). Approaches to studying when preparing for final exams as a function of coping strategies. *Personality and Individual Differences, 43*(1), 191-202. doi:10.1016/j.paid.2006.12.002

Moran, T. P. (2016). Anxiety and Working Memory Capacity: A Meta-Analysis and Narrative Review. *Psychological Bulletin, 142*(8), 831-864. doi:10.1037/bul0000051

Morris, L. W., Davis, M. A., & Hutchings, C. H. (1981). COGNITIVE AND EMOTIONAL COMPONENTS OF ANXIETY - LITERATURE-REVIEW AND A REVISED WORRY-EMOTIONALITY SCALE. *Journal of Educational Psychology, 73*(4), 541-555. doi:10.1037/0022-0663.73.4.541

Morris, L. W., & Liebert, R. M. (1970). RELATIONSHIP OF COGNITIVE AND EMOTIONAL COMPONENTS OF TEST ANXIETY TO PHYSIOLOGICAL AROUSAL AND ACADEMIC PERFORMANCE. *Journal of Consulting and Clinical Psychology, 35*(3), 332-&. doi:10.1037/h0030132

Mosca, O., Lauriola, M., & Carleton, R. N. (2016). Intolerance of Uncertainty: A Temporary Experimental Induction Procedure. *Plos One, 11*(6), e0155130. doi:10.1371/journal.pone.0155130

Mowrer, O. H. (1947). On the dual nature of learniná a re-interpretation of "conditioning" and "problem-solving.". *Harvard Educational Review*.

Mueller, J. H. (1992). TEST ANXIETY AND IMPLICIT MEMORY. *International Journal of Psychology, 27*(3-4), 124-124. Retrieved from <Go to ISI>://WOS:A1992JF42000650

Mueller, J. H., & Courtois, M. R. (1980). TEST ANXIETY AND BREADTH OF ENCODING EXPERIENCES IN FREE-RECALL. *Journal of Research in Personality, 14*(4), 458-466. doi:10.1016/0092-6566(80)90004-5

Mueller, J. H., Elser, M. J., & Rollack, D. N. (1993). TEST ANXIETY AND IMPLICIT MEMORY. *Bulletin of the Psychonomic Society, 31*(6), 531-533. doi:10.3758/bf03337344

Muis, K. R., Ranellucci, J., Franco, G. M., & Crippen, K. J. (2013). The Interactive Effects of Personal Achievement Goals and Performance Feedback in an Undergraduate Science Class. *Journal of Experimental Education, 81*(4), 556-578. doi:10.1080/00220973.2012.738257

Muthén, L. K., & Muthén, B. O. (2008-2017). *Mplus User's Guide*. Los Angeles, CA: Muthén & Muthén.

Naveh-Benjamin, M., Lavi, H., McKeachie, W. J., & Lin, Y. G. (1997). Individual differences in students' retention of knowledge and conceptual structures learned in University and High school courses: The case of test anxiety. *Applied Cognitive Psychology, 11*(6), 507-526. doi:10.1002/(sici)1099-0720(199712)11:6<507::Aid-acp482>3.0.Co;2-g

Ne'Eman-Haviv, V., & Bonny-Noach, H. (2019). Substances as Self-Treatment for Cognitive Test Anxiety among Undergraduate Students. *J Psychoactive Drugs, 51*(1), 78-84. doi:10.1080/02791072.2018.1564090

Nelson, T. O., & Narens, L. (1990). Metamemory: A Theoretical Framework and New Findings. In G. H. Bower (Ed.), *Psychology of Learning and Motivation* (Vol. 26, pp. 125-173): Academic Press.

Neuderth, S., Jabs, B., & Schmidtke, A. (2009). Strategies for reducing test anxiety and optimizing exam preparation in German university students: a prevention-oriented pilot project of the University of Wurzburg. *Journal of Neural Transmission, 116*(6), 785-790. doi:10.1007/s00702-008-0123-7

Newble, D. (2004). Techniques for measuring clinical competence: objective structured clinical examinations. *Medical Education, 38*(2), 199-203. doi:10.1046/j.1365-2923.2004.01755.x

Newman, M. G., Jacobson, N. C., Zainal, N. H., Shin, K. E., Szkodny, L. E., & Sliwinski, M. J. (2019). The effects of worry in daily life: An ecological momentary assessment study supporting the tenets of the contrast avoidance model. *Clinical Psychological Science, 7*(4), 794-810. doi:10.1177/2167702619827019

Newman, M. G., & Llera, S. J. (2011). A novel theory of experiential avoidance in generalized anxiety disorder: a review and synthesis of research supporting a contrast avoidance model of worry. *Clinical Psychology Review, 31*(3), 371-382. doi:10.1016/j.cpr.2011.01.008

Nicholson, B., & Forrest, K. (2009). What influences performance in the OSCE exam? The medical student perspective. *Medical Teacher, 31*(11), 1040-1041. Retrieved from <Go to ISI>://WOS:000273159000019

Nicholson, L., & Putwain, D. (2019). 'We should just be told to try our best'. *Psychologist, 32*, 39-41. Retrieved from <Go to ISI>://WOS:000465593000022

Normann, N., & Morina, N. (2018). The Efficacy of Metacognitive Therapy: A Systematic Review and Meta-Analysis. *Frontiers in Psychology, 9*(2211). doi:10.3389/fpsyg.2018.02211

Norton, P. J. (2005). A psychometric analysis of the Intolerance of Uncertainty Scale among four racial groups. *Journal of Anxiety Disorders, 19*(6), 699-707. doi:10.1016/j.janxdis.2004.08.002

Novick-Kline, P., Turk, C. L., Mennin, D. S., Hoyt, E. A., & Gallagher, C. L. (2005). Level of emotional awareness as a differentiating variable between individuals with and without generalized anxiety disorder. *J Anxiety Disord, 19*(5), 557-572. doi:10.1016/j.janxdis.2004.06.001

O'Carroll, P. J., & Fisher, P. (2013). Metacognitions, worry and attentional control in predicting OSCE performance test anxiety. *Medical Education, 47*(6), 562-568. doi:10.1111/medu.12125

O'Cathain, A., Croot, L., Duncan, E., Rousseau, N., Sworn, K., Turner, K. M., . . . Hoddinott, P. (2019). Guidance on how to develop complex interventions to improve health and healthcare. *BMJ Open, 9*(8), e029954. doi:10.1136/bmjopen-2019-029954

O'Neil, H. F., & Brown, R. S. (1998). Differential effects of question formats in math assessment on metacognition and affect. *Applied Measurement in Education, 11*(4), 331-351. doi:10.1207/s15324818ame1104_3

Oglesby, M. E., Allan, N. P., Short, N. A., Raines, A. M., & Schmidt, N. B. (2017). Factor Mixture Modeling of Intolerance of Uncertainty. *Psychological Assessment, 29*(4), 435-445. doi:10.1037/pas0000357

Olsson, F. M. (2008). *New developments in the psychology of motivation*. New York: Nova Biomedical Books.

Orbach, G., Lindsay, S., & Grey, S. (2007). A randomised placebo-controlled trial of a self-help Internet-based intervention for test anxiety. *Behaviour Research and Therapy, 45*(3), 483-496. doi:10.1016/j.brat.2006.04.002

Ost, L. G. (2008). Efficacy of the third wave of behavioral therapies: A systematic review and meta-analysis. *Behaviour Research and Therapy, 46*(3), 296-321. doi:10.1016/j.brat.2007.12.005

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., . . . Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ, 372*, n71. doi:10.1136/bmj.n71

Papageorgiou, C., & Wells, A. (2003). An Empirical Test of a Clinical Metacognitive Model of Rumination and Depression. *Cognitive Therapy and Research, 27*(3), 261-273. doi:10.1023/A:1023962332399

Papageorgiou, C., & Wells, A. (2009). A prospective test of the clinical metacognitive model of rumination and depression. *International Journal of Cognitive Therapy, 2*(2), 123-131. doi:10.1521/ijct.2009.2.2.123

Pate, A. N., Neely, S., Malcom, D. R., Daugherty, K. K., Zagar, M., & Medina, M. S. (2021). Multisite Study Assessing the Effect of Cognitive Test Anxiety on Academic and Standardized Test Performance. *American Journal of Pharmaceutical Education, 85*(1), 8041-8041. doi:10.5688/ajpe8041

Paulman, R. G., & Kennelly, K. J. (1984). Test anxiety and ineffective test taking: Different names, same construct? *Journal of Educational Psychology, 76*(2), 279-288. doi:10.1037/0022-0663.76.2.279

Payne, B. D. (1984). THE RELATIONSHIP OF TEST ANXIETY AND ANSWER-CHANGING BEHAVIOR - AN ANALYSIS BY RACE AND SEX. *Measurement and Evaluation in Guidance, 16*(4), 205-210. doi:10.1080/00256307.1984.12022358

Payne, B. D., Smith, J. E., & Payne, D. A. (1983a). GRADE, SEX, AND RACE DIFFERENCES IN TEST ANXIETY. *Psychological Reports, 53*(1), 291-294. doi:10.2466/pr0.1983.53.1.291

Payne, B. D., Smith, J. E., & Payne, D. A. (1983b). SEX AND ETHNIC-DIFFERENCES IN RELATIONSHIPS OF TEST ANXIETY TO PERFORMANCE IN SCIENCE EXAMINATIONS BY 4TH + 8TH GRADE STUDENTS - IMPLICATIONS FOR VALID

INTERPRETATIONS OF ACHIEVEMENT-TEST SCORES. *Educational and Psychological Measurement, 43*(1), 267-270. doi:10.1177/001316448304300135

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. doi:10.1007/s10648-006-9029-9

Pekrun, R., Goetz, T., Perry, R. P., Kramer, K., Hochstadt, M., & Molfenter, S. (2004). Beyond test anxiety: Development and vailidation of the Test Emotions Questionnaire (TEQ). *Anxiety Stress and Coping, 17*(3), 287-316. doi:10.1080/10615800412331303847

Penney, A. M., Rachor, G. S., & Deleurme, K. A. (2020). Differentiating the roles of intolerance of uncertainty and negative beliefs about worry across emotional disorders. *Journal of Experimental Psychopathology, 11*(4), 2043808720970072. doi:10.1177/2043808720970072

Pepperdine, E., Lomax, C., & Freeston, M. H. (2018). Disentangling intolerance of uncertainty and threat appraisal in everyday situations. *Journal of Anxiety Disorders, 57*, 31-38. doi:10.1016/j.jamcdis.2018.04.002

Pintrich, P. R. (2000). Multiple goals, multiple pathways: The role of goal orientation in learning and achievement. *Journal of Educational Psychology, 92*(3), 544-555. doi:10.1037/0022-0663.92.3.544

Pintrich, P. R., Smith, D. A. F., Garcia, T., & McKeachie, W. J. (1993). RELIABILITY AND PREDICTIVE-VALIDITY OF THE MOTIVATED STRATEGIES FOR LEARNING QUESTIONNAIRE

(MSLQ). *Educational and Psychological Measurement, 53*(3), 801-813. doi:10.1177/0013164493053003024

Pintrich, P. R., & Zusho, A. (2007). *STUDENT MOTIVATION AND SELF-REGULATED LEARNING IN THE COLLEGE CLASSROOM*.

Prochaska, J. O. (1971). Symptom and dynamic cues in the implosive treatment of test anxiety *Journal of Abnormal Psychology, 77*(2), 133-+. doi:10.1037/h0030738

Purdon, C., Rowa, K., & Antony, M. M. (2005). Thought suppression and its effects on thought frequency, appraisal and mood state in individuals with obsessive-compulsive disorder. *Behaviour Research and Therapy, 43*(1), 93-108. doi:https://doi.org/10.1016/j.brat.2003.11.007

Putwain, D. (2008). Examination stress and test anxiety. *Psychologist, 21*(12), 1026-1029. Retrieved from <Go to ISI>://WOS:000261732400020

Putwain, D., & Symes, W. (2011). Perceived fear appeals and examination performance: Facilitating or debilitating outcomes? *Learning and Individual Differences, 21*(2), 227-232. doi:10.1016/j.lindif.2010.11.022

Putwain, D. W. (2007). Test anxiety in UK schoolchildren: Prevalence and demographic patterns. *British Journal of Educational Psychology, 77*, 579-593. doi:10.1348/000709906x161704

Putwain, D. W. (2019). An examination of the self-referent executive processing model of test anxiety: control, emotional regulation, self-handicapping, and examination performance. *European Journal of Psychology of Education, 34*(2), 341-358. doi:10.1007/s10212-018-0383-z

Putwain, D. W., Daly, A. L., Chamberlain, S., & Sadreddini, S. (2016). ""Sink or swim': buoyancy and coping in the cognitive test anxiety - academic

performance relationship". *Educational Psychology, 36*(10), 1807-1825. doi:10.1080/01443410.2015.1066493

Putwain, D. W., Langdale, H. C., Woods, K. A., & Nicholson, L. J. (2011). Developing and piloting a dot-probe measure of attentional bias for test anxiety. *Learning and Individual Differences, 21*(4), 478-482. doi:10.1016/j.lindif.2011.02.002

Putwain, D. W., & Symes, W. (2012). Achievement goals as mediators of the relationship between competence beliefs and test anxiety. *British Journal of Educational Psychology, 82*(2), 207-224. doi:10.1111/j.2044-8279.2011.02021.x

Putwain, D. W., Symes, W., Coxon, E., & Gallard, D. (2020). Attention bias in test anxiety: the impact of a test-threat congruent situation, presentation time, and approach-avoidance temperament. *Educational Psychology, 40*(6), 713-734. doi:10.1080/01443410.2020.1740653

Putwain, D. W., von der Embse, N. P., Rainbird, E. C., & West, G. (2021). The development and validation of a new Multidimensional Test Anxiety Scale (MTAS). *European Journal of Psychological Assessment, 37*(3), 236-246. doi:10.1027/1015-5759/a000604

Rajiah, K., & Saravanan, C. (2014). The effectiveness of psychoeducation and systematic desensitization to reduce test anxiety among first-year pharmacy students. *American Journal of Pharmaceutical Education, 78*(9). Retrieved from <Go to ISI>://WOS:000346349400006

Ree, M. J., French, D., MacLeod, C., & Locke, V. (2008). Distinguishing Cognitive and Somatic Dimensions of State and Trait Anxiety: Development and Validation of the State-Trait Inventory for Cognitive

and Somatic Anxiety (STICSA). *Behavioural and Cognitive Psychotherapy, 36*(3), 313-332. doi:10.1017/s1352465808004232

Reed, M., & Saslow, C. (1980). The effects of relaxation instructions and EMG biofeedback on test anxiety, general anxiety, and locus of control *Journal of Clinical Psychology, 36*(3), 683-690. doi:10.1002/1097-4679(198007)36:3<683::aid-jclp2270360313>3.0.co;2-p

Reed, R., & Meyer, R. G. (1974). REDUCTION OF TEST ANXIETY VIA AUTOGENIC THERAPY. *Psychological Reports, 35*(1), 649-650. Retrieved from <Go to ISI>://WOS:A1974U631100072

Regehr, C., Glancy, D., & Pitts, A. (2013). Interventions to reduce stress in university students: A review and meta-analysis. *Journal of Affective Disorders, 148*(1), 1-11. doi:10.1016/j.jad.2012.11.026

Reiss, N., Warnecke, I., Tolgou, T., Krampen, D., Luka-Krausgrill, U., & Rohrmann, S. (2017). Effects of cognitive behavioral therapy with relaxation vs. imagery rescripting on test anxiety: A randomized controlled trial. *Journal of Affective Disorders, 208*, 483-489. doi:10.1016/j.jad.2016.10.039

Richardson, M., Abraham, C., & Bond, R. (2012). Psychological Correlates of University Students' Academic Performance: A Systematic Review and Meta-Analysis. *Psychological Bulletin, 138*(2), 353-387. doi:10.1037/a0026838

Ricketts, M. S., & Galloway, R. E. (1984). Effects of three different one-hour single-session treatments for test anxiety *Psychological Reports, 54*(1), 115-120. Retrieved from <Go to ISI>://WOS:A1984SL20800024

Roberts, K. E., Hart, T. A., & Eastwood, J. D. (2016). Factor Structure and Validity of the State-Trait Inventory for Cognitive and Somatic Anxiety. *Psychological Assessment, 28*(2), 134-146. doi:10.1037/pas0000155

Robichaud, M., & Dugas, M. J. (2005a). Negative problem orientation (Part I): psychometric properties of a new measure. *Behaviour Research and Therapy, 43*(3), 391-401. doi:https://doi.org/10.1016/j.brat.2004.02.007

Robichaud, M., & Dugas, M. J. (2005b). Negative problem orientation (Part II): Construct validity and specificity to worry. *Behaviour Research and Therapy, 43*(3), 403-412. doi:10.1016/j.brat.2004.02.008

Rodway, C., Tham, S. G., Ibrahim, S., Turnbull, P., Windfuhr, K., Shaw, J., . . . Appleby, L. (2016). Suicide in children and young people in England: a consecutive case series. *Lancet Psychiatry, 3*(8), 751-759. doi:10.1016/s2215-0366(16)30094-3

Rogaten, J., & Moneta, G. B. (2017). Positive and Negative Structures and Processes Underlying Academic Performance: A Chained Mediation Model. *Journal of Happiness Studies, 18*(4), 1095-1119. doi:10.1007/s10902-016-9765-6

Rogaten, J., Moneta, G. B., & Spada, M. M. (2013). Academic Performance as a Function of Approaches to Studying and Affect in Studying. *Journal of Happiness Studies, 14*(6), 1751-1763. doi:10.1007/s10902-012-9408-5

Rogers, C. (1961). *On Becoming a Person: A Therapist's View of Psychotherapy*. London: Constable.

Roma, V. G., & Hope, D. A. (2017). Are we certain about which measure of intolerance of uncertainty to use yet? *Cognitive Behaviour Therapy, 46*(4), 327-338. doi:10.1080/16506073.2016.1252792

Romano, J. L., & Cabianca, W. A. (1978). EMG biofeedback training versus systematic desensitization for test anxiety reduction *Journal of Counseling Psychology, 25*(1), 8-13. Retrieved from <Go to ISI>://WOS:A1978EJ18300002

Roos, A. L., Goetz, T., Krannich, M., Jarrell, A., Donker, M., & Mainhard, T. (2021). Test anxiety components: an intra-individual approach testing their control antecedents and effects on performance. *Anxiety Stress and Coping, 34*(3), 279-298. doi:10.1080/10615806.2020.1850700

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. doi:10.1007/s10648-020-09543-z

Rosenthal, R. (1979). The file drawer problem and tolerance for null results *Psychological Bulletin, 86*(3), 638-641. doi:10.1037//0033-2909.86.3.638

Ross, E. (1968). Effects of challenging and supportive instructions on verbal learning in older persons. *Journal of Educational Psychology, 59*, 261–266.

Rosser, B. A. (2019). Intolerance of Uncertainty as a Transdiagnostic Mechanism of Psychological Difficulties: A Systematic Review of Evidence Pertaining to Causality and Temporal Precedence. *Cognitive Therapy and Research, 43*(2), 438-463. doi:10.1007/s10608-018-9964-z

Roussis, P., & Wells, A. (2006). Post-traumatic stress symptoms: Tests of relationships with thought control strategies and beliefs as predicted by the metacognitive model. *Personality and Individual Differences, 40*(1), 111-122. doi:10.1016/j.paid.2005.06.019

Rückert, H.-W. (2015). Students′ mental health and psychological counselling in Europe. *Mental Health & Prevention, 3*(1), 34-40. doi:https://doi.org/10.1016/j.mhp.2015.04.006

Russell, R. K. (1981). TEST ANXIETY - THEORY, RESEARCH, AND APPLICATIONS - SARASON,IG. *Contemporary Psychology, 26*(6), 429-430. Retrieved from <Go to ISI>://WOS:A1981LU71900006

Russell, R. K., & Lent, R. W. (1982). Cue-controlled relaxation and systematic desensitization versus nonspecific factors in treating test anxiety *Journal of Counseling Psychology, 29*(1), 100-103. Retrieved from <Go to ISI>://WOS:A1982NE21000011

Russell, R. K., Miller, D. E., & June, L. N. (1974). GROUP CUE-CONTROLLED RELAXATION IN TREATMENT OF TEST ANXIETY. *Behavior Therapy, 5*(4), 572-573. doi:10.1016/s0005-7894(74)80050-x

Russell, R. K., Miller, D. E., & June, L. N. (1975). COMPARISON BETWEEN GROUP SYSTEMATIC DESENSITIZATION AND CUE-CONTROLLED RELAXATION IN TREATMENT OF TEST ANXIETY. *Behavior Therapy, 6*(2), 172-177. doi:10.1016/s0005-7894(75)80138-9

Russell, R. K., & Sipich, J. F. (1973). CUE-CONTROLLED RELAXATION IN TREATMENT OF TEST ANXIETY. *Journal of Behavior Therapy and Experimental Psychiatry, 4*(1), 47-49. doi:10.1016/0005-7916(73)90038-4

Russell, R. K., Wise, F., & Stratoudakis, J. P. (1976). Treatment of test anxiety by cue-controlled relaxation and systematic desensitization *Journal of*

Counseling Psychology, *23*(6), 563-566. doi:10.1037//0022-0167.23.6.563

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions. *Contemporary Educational Psychology*, *25*(1), 54-67. doi:https://doi.org/10.1006/ceps.1999.1020

Sapp, M. (1996). Three treatments for reducing the worry and emotionality components of test anxiety with undergraduate and graduate college students: Cognitive-behavioral hypnosis, relaxation therapy, and supportive counseling. *Journal of College Student Development*, *37*(1), 79-87. Retrieved from <Go to ISI>://WOS:A1996TP32400009

Sarason, I. G. (1978). The Test Anxiety Scale: Concept and research. In C. D. S. I. G. Sarason & (Eds.) (Eds.), *Stress and anxiety (Vol. 5, pp. 193–216)*. Washington, DC: Hemisphere.

Sarason, I. G. (1980). Introduction to the Study of Test Anxiety. In I. G. Sarason (Ed.), *Test Anxiety: Theory, Research, and Applications* (pp. 3-14). Hillsdale, NJ: Lawrence Erlbaum Associates.

Sarason, I. G. (1981). TEST ANXIETY, STRESS, AND SOCIAL SUPPORT. *Journal of Personality*, *49*(1), 101-114. doi:10.1111/j.1467-6494.1981.tb00849.x

Sarason, I. G. (1984). STRESS, ANXIETY, AND COGNITIVE INTERFERENCE - REACTIONS TO TESTS. *Journal of Personality and Social Psychology*, *46*(4), 929-938. doi:10.1037/0022-3514.46.4.929

Sarason, I. G. (1988). Anxiety, self-preoccupation and attention. *Anxiety Research*, *1*(1), 3–7.

Sarason, I. G., & Sarason, B. R. (1990). Test anxiety. In *Handbook of social and evaluation anxiety.* (pp. 475-495). New York, NY, US: Plenum Press.

Sarason, I. G., Sarason, B. R., & Pierce, G. R. (1990). ANXIETY, COGNITIVE INTERFERENCE, AND PERFORMANCE. *Journal of Social Behavior and Personality, 5*(2), 1-18. Retrieved from <Go to ISI>://WOS:A1990CX04000001

Sarason, I. G., & Stoops, R. (1978). TEST ANXIETY AND PASSAGE OF TIME. *Journal of Consulting and Clinical Psychology, 46*(1), 102-109. doi:10.1037/0022-006x.46.1.102

Sarason, S. B., & Mandler, G. (1952). SOME CORRELATES OF TEST ANXIETY. *Journal of Abnormal and Social Psychology, 47*(4), 810-817. doi:10.1037/h0060009

Saravanan, C., & Kingston, R. (2014). A randomized control study of psychological intervention to reduce anxiety, amotivation and psychological distress among medical students. *Journal of Research in Medical Sciences, 19*(5), 391-397. Retrieved from <Go to ISI>://WOS:000342790000004

Saravanan, C., Kingston, R., & Gin, M. (2014). Is test anxiety a problem among medical students: A cross sectional study on outcome of test anxiety among medical students? *International Journal of Psychological Studies, 6*(3), 24-31.

Sattler, S. (2019). Nonmedical use of prescription drugs for cognitive enhancement as response to chronic stress especially when social support is lacking. *Stress and Health, 35*(2), 127-137. doi:10.1002/smi.2846

Sattler, S., & Wiegel, C. (2013). Cognitive test anxiety and cognitive enhancement: the influence of students' worries on their use of performance-enhancing drugs. *Subst Use Misuse, 48*(3), 220-232. doi:10.3109/10826084.2012.751426

Schaefer, A., Mattheg, H., Pfitzer, G., & Koehle, K. (2007). Mental health and performance of medical students with high and low test anxiety. *Psychotherapie Psychosomatik Medizinische Psychologie, 57*(7), 289-297. doi:10.1055/s-2006-951974

Schulz, K. F., Altman, D. G., Moher, D., & Grp, C. (2010). CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. *Bmc Medicine, 8*. doi:10.1186/1741-7015-8-18

Schwarzer, R. (1981). THE IDENTIFICATION OF WORRY AND OF EMOTIONALITY FACTORS IN TEST ANXIETY. *Psychologische Beitrage, 23*(3-4), 579-594. Retrieved from <Go to ISI>://WOS:A1981NJ88500018

Schwarzer, R. (1984). WORRY AND EMOTIONALITY AS SEPARATE COMPONENTS IN TEST ANXIETY. *International Review of Applied Psychology-Revue Internationale De Psychologie Appliquee, 33*(2), 205-220. Retrieved from <Go to ISI>://WOS:A1984ST48000004

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. doi:10.1002/pits.21689

Segool, N. K., von der Embse, N. P., Mata, A. D., & Gallant, J. (2014). Cognitive Behavioral Model of Test Anxiety in a High-Stakes Context: An Exploratory Study. *School Mental Health, 6*(1), 50-61. doi:10.1007/s12310-013-9111-7

Seipp, B. (1991). Anxiety and academic performance: A meta-analysis of findings. *Anxiety Research, 4*(1), 27-41. doi:10.1080/08917779108248762

Seipp, B., & Schwarzer, C. (1996). Cross-cultural anxiety research: A review. In C. Schwarzer & M. Zeidner (Eds.), *Stress, anxiety, and coping in academic settings* (pp. 13-68). Tubingen, Germany: Francke-Verlag.

Sellers, R., Varese, F., Wells, A., & Morrison, A. P. (2017). A meta-analysis of metacognitive beliefs as implicated in the self-regulatory executive function model in clinical psychosis. *Schizophr Res, 179*, 75-84. doi:10.1016/j.schres.2016.09.032

Sexton, K. A., & Dugas, M. J. (2009). Defining Distinct Negative Beliefs About Uncertainty: Validating the Factor Structure of the Intolerance of Uncertainty Scale. *Psychological Assessment, 21*(2), 176-186. doi:10.1037/a0015827

Shi, Z., Gao, X., & Zhou, R. L. (2014). Emotional working memory capacity in test anxiety. *Learning and Individual Differences, 32*, 178-183. doi:10.1016/j.lindif.2014.03.011

Shihata, S., McEvoy, P. M., & Mullan, B. A. (2018). A Bifactor Model of Intolerance of Uncertainty in Undergraduate and Clinical Samples: Do We Need to Reconsider the Two-Factor Model? *Psychological Assessment, 30*(7), 893-903. doi:10.1037/pas0000540

Shihata, S., McEvoy, P. M., Mullan, B. A., & Carleton, R. N. (2016). Intolerance of uncertainty in emotional disorders: What uncertainties remain? *Journal of Anxiety Disorders, 41*, 115-124. doi:10.1016/j.janxdis.2016.05.001

Sibrava, N. J., & Borkovec, T. D. (2006). The cognitive avoidance theory of worry. In G. C. L. Davey & A. Wells (Eds.), *Worry and its Psychological Disorders: Theory, Assessment and Treatment* (pp. 239-258). Chichester, UK: Wiley.

Skinner, B. F. (1938). *The behavior of organisms: an experimental analysis*. Oxford, England: Appleton-Century.

Smith, M. L., & Glass, G. V. (1977). Meta-analysis of psychotherapy outcome studies. *Am Psychol, 32*(9), 752-760. doi:10.1037//0003-066x.32.9.752

Smith, T. W., Snyder, C. R., & Handelsman, M. M. (1982). On the self-serving function of an academic wooden leg: Test anxiety as a self-handicapping strategy. *Journal of Personality and Social Psychology, 42*(2), 314-321. doi:10.1037/0022-3514.42.2.314

Snyder, A. L., & Deffenbacher, J. L. (1977). COMPARISON OF RELAXATION AS SELF-CONTROL AND SYSTEMATIC-DESENSITIZATION IN TREATMENT OF TEST ANXIETY. *Journal of Consulting and Clinical Psychology, 45*(6), 1202-1203. doi:10.1037/0022-006x.45.6.1202

Sommer, M., & Arendasy, M. E. (2014). Comparing different explanations of the effect of test anxiety on respondents' test scores. *Intelligence, 42*, 115-127. doi:10.1016/j.intell.2013.11.003

Spada, M. M., Georgiou, G. A., & Wells, A. (2010). The Relationship among Metacognitions, Attentional Control, and State Anxiety. *Cognitive Behaviour Therapy, 39*(1), 64-71. doi:10.1080/16506070902991791

Spada, M. M., Mohiyeddini, C., & Wells, A. (2008). Measuring metacognitions associated with emotional distress: Factor structure and predictive validity of the metacognitions questionnaire 30. *Personality and Individual Differences, 45*(3), 238-242. doi:10.1016/j.paid.2008.04.005

Spada, M. M., & Moneta, G. B. (2012). A metacognitive-motivational model of surface approach to studying. *Educational Psychology, 32*(1), 45-62. doi:10.1080/01443410.2011.625610

Spada, M. M., & Moneta, G. B. (2014). Metacognitive and motivational predictors of surface approach to studying and academic examination performance. *Educational Psychology, 34*(4), 512-523. doi:10.1080/01443410.2013.814196

Spada, M. M., Nikcevic, A. V., Moneta, G. B., & Ireson, J. (2006). Metacognition as a Mediator of the Effect of Test Anxiety on a Surface Approach to Studying. *Educational Psychology, 26*(5), 615-624. doi:10.1080/01443410500390673

Spiegler, M. D., Morris, L. W., & Liebert, R. M. (1968). COGNITIVE AND EMOTIONAL COMPONENTS OF TEST ANXIETY - TEMPORAL FACTORS. *Psychological Reports, 22*(2), 451-&. doi:10.2466/pr0.1968.22.2.451

Spielberger, C. D. (1966). *Theory and research on anxietyl*. New York: Academic Press.

Spielberger, C. D. (1980). *Test anxiety inventory: Preliminary professional manual*. Palo Alto, CA: Consulting Psychologist Press.

Spielberger, C. D., Anton, W. D., & Bedell, J. (1976). The nature and treatment of test anxiety. In M. Zuckerman & C. D. Spielberger (Eds.), *Emotions and anxiety: New concepts, methods, and applications*. New Jersey: Lawrence Erlbaum Associates.

Spielberger, C. D., Gorsuch, R. L., Lushene, R., Vagg, P. R., & Jacobs, G. A. (1983). *Manual for the State-Trait Anxiety Inventory*. Palo Alto, CA: Consulting Psychologists Press.

Spielberger, C. D., & Vagg, P. R. (1995). Test anxiety: A transactional process model. In *Test anxiety: Theory, assessment, and treatment.* (pp. 3-14). Philadelphia, PA, US: Taylor & Francis.

Spielberger, C. D., & Vagg, P. R. (1995). *Test anxiety: Theory, assessment, and treatment*. London: Taylor & Francis Ltd.

Steel, P. (2007). The nature of procrastination: A meta-analytic and theoretical review of quintessential self-regulatory failure. *Psychological Bulletin, 133*(1), 65-94. doi:10.1037/0033-2909.133.1.65

Stiles, W. B., Barkham, M., & Wheeler, S. (2015). Duration of psychological therapy: Relation to recovery and improvement rates in UK routine practice. *British Journal of Psychiatry, 207*(2), 115-122. doi:10.1192/bjp.bp.114.145565

Stober, J. (2004). Dimensions of test anxiety: Relations to ways of coping with pre-exam anxiety and uncertainty. *Anxiety Stress and Coping, 17*(3), 213-226. doi:10.1080/10615800412331292615

Stober, J., & Pekrun, R. (2004). Advances in test anxiety research. *Anxiety Stress and Coping, 17*(3), 205-211. doi:10.1080/1061580412331303225

Sun, X., Zhu, C., & So, S. H. W. (2017). Dysfunctional metacognition across psychopathologies: A meta-analytic review. *European Psychiatry, 45*, 139-153. doi:10.1016/j.eurpsy.2017.05.029

Szafranski, D. D., Barrera, T. L., & Norton, P. J. (2012). Test anxiety inventory: 30 years later. *Anxiety Stress and Coping, 25*(6), 667-677. doi:10.1080/10615806.2012.663490

Team, R. C. (2013). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Tektaş, O. Y., Paulsen, F., & Sel, S. (2013). Test anxiety among German medical students and its impact on lifestyle and substance abuse. *Med Teach, 35*(11), 969. doi:10.3109/0142159x.2013.786813

Tempel, T., & Neumann, R. (2016). Taming Test Anxiety: The Activation of Failure-Related Concepts Enhances Cognitive Test Performance of Test-Anxious Students. *Journal of Experimental Education, 84*(4), 702-722. doi:10.1080/00220973.2015.1094649

Temple, J., Salmon, P., Tudur-Smith, C., Huntley, C. D., & Fisher, P. L. (2018). A systematic review of the quality of randomized controlled trials of psychological treatments for emotional distress in breast cancer. *Journal of Psychosomatic Research, 108*, 22-31. doi:10.1016/j.jpsychores.2018.02.013

Thames, A. D., Panos, S. E., Arentoft, A., Byrd, D. A., Hinkin, C. H., & Arbid, N. (2015). Mild Test Anxiety Influences Neurocognitive Performance Among African Americans and European Americans: Identifying

Interfering and Facilitating Sources. *Cultural Diversity & Ethnic Minority Psychology, 21*(1), 105-113. doi:10.1037/a0037530

Thomas, C. L., & Cassady, J. C. (2019). The influence of personality factors, value appraisals, and control appraisals on cognitive test anxiety. *Psychology in the Schools, 56*(10), 1568-1582. doi:10.1002/pits.22303

Thomas, C. L., Cassady, J. C., & Finch, W. H. (2018). Identifying Severity Standards on the Cognitive Test Anxiety Scale: Cut Score Determination Using Latent Class and Cluster Analysis. *Journal of Psychoeducational Assessment, 36*(5), 492-508. doi:10.1177/0734282916686004

Thomas, C. L., Cassady, J. C., & Heller, M. L. (2017). The influence of emotional intelligence, cognitive test anxiety, and coping strategies on undergraduate academic performance. *Learning and Individual Differences, 55*, 40-48. doi:10.1016/j.lindif.2017.03.001

Tibbett, T. P., & Ferrari, J. R. (2019). Return to the origin: what creates a procrastination identity? *Current Issues in Personality Psychology, 7*(1), 1-7. doi:10.5114/cipp.2018.75648

Tobias, S. (1985). Test Anxiety: Interference, Defective Skills, and Cognitive Capacity. *Educational Psychologist, 20*(3), 135-142. doi:10.1207/s15326985ep2003_3

Tobias, S. (1992). The impact of test anxiety on cognition in school learning. In K. A. Hagtvet & B. T. Johnsen (Eds.), *Advances in test anxiety research (Vol. 7)* (pp. 18-31). Lisse, The Netherlands: Swets & Zeitlinger.

Tobias, S., & Everson, H. T. (1997). Studying the relationship between affective and metacognitive variables. *Anxiety Stress and Coping, 10*(1), 59-81. doi:10.1080/10615809708249295

Tolin, D. F., Abramowitz, J. S., Brigidi, B. D., & Foa, E. B. (2003). Intolerance of uncertainty in obsessive-compulsive disorder. *J Anxiety Disord, 17*(2), 233-242. doi:10.1016/s0887-6185(02)00182-2

Tosevski, D. L., Milovancevic, M. P., & Gajic, S. D. (2010). Personality and psychopathology of university students. *Curr Opin Psychiatry, 23*(1), 48-52. doi:10.1097/YCO.0b013e328333d625

Tsegay, L., Shumet, S., Damene, W., Gebreegziabhier, G., & Ayano, G. (2019). Prevalence and determinants of test anxiety among medical students in Addis Ababa Ethiopia. *BMC Medical Education, 19*(1), 423. doi:10.1186/s12909-019-1859-5

van der Heiden, C., Muris, P., & van der Molen, H. T. (2012). Randomized controlled trial on the effectiveness of metacognitive therapy and intolerance-of-uncertainty therapy for generalized anxiety disorder. *Behaviour Research and Therapy, 50*(2), 100-109. doi:https://doi.org/10.1016/j.brat.2011.12.005

van Eerde, W. (2003). A meta-analytically derived nomological network of procrastination. *Personality and Individual Differences, 35*(6), 1401-1418. doi:10.1016/s0191-8869(02)00358-6

von der Embse, N. P., Putwain, D. W., & Francis, G. (2021). Interpretation and use of the Multidimensional Test Anxiety Scale (MTAS). *School Psychology, 36*(2), 86-96. doi:10.1037/spq0000427

Viechtbauer, W. (2010). Conducting Meta-Analyses in R with the metafor Package. *Journal of Statistical Software, 36*(3), 1-48. Retrieved from <Go to ISI>://WOS:000281593200001

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. doi:10.1016/j.jad.2017.11.048

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. doi:10.1177/0734282913504541

von der Embse, N. P., Putwain, D. W., & Francis, G. (2021). Interpretation and Use of the Multidimensional Test Anxiety Scale (MTAS). *School Psychology, 36*(2), 86-96. doi:10.1037/spq0000427

Watson, J. B. (1913). Psychology as the behaviorist views it. *Psychological Review, 20*(2), 158-177. doi:10.1037/h0074428

Wegner, D. M., Schneider, D. J., Carter, S. R., & White, T. L. (1987). Paradoxical effects of thought suppression. *Journal of Personality and Social Psychology, 53*(1), 5-13. doi:10.1037/0022-3514.53.1.5

Wells, A. (2000). *Emotional disorders and metacognition: Innovative cognitive therapy*. Chichester, UK: Wiley.

Wells, A. (2005). The metacognitive model of GAD: Assessment of meta-worry and relationship with DSM-IV generalized anxiety disorder. *Cognitive Therapy and Research, 29*(1), 107-121. doi:10.1007/s10608-005-1652-0

Wells, A. (2009). *Metacognitive therapy for anxiety and depression*. London: The Guildford Press.

Wells, A., & Cartwright-Hatton, S. (2004). A short form of the metacognitions questionnaire: properties of the MCQ-30. *Behaviour Research and Therapy, 42*(4), 385-396. doi:10.1016/s0005-7967(03)00147-5

Wells, A., & Matthews, G. (1994). *Attention and emotion: A clinical perspective*. Hove, UK: Erlbaum.

Wells, A., & Matthews, G. (1996a). Anxiety and cognition. *Current Opinion in Psychiatry, 9*(6), 422-426. doi:10.1097/00001504-199611000-00011

Wells, A., & Matthews, G. (1996b). Modelling cognition in emotional disorder: The S-REF model. *Behaviour Research and Therapy, 34*(11), 881-888. doi:https://doi.org/10.1016/S0005-7967(96)00050-2

Wells, A., & Papageorgiou, C. (1998). Relationships between worry, obsessive-compulsive symptoms and meta-cognitive beliefs. *Behaviour Research and Therapy, 36*(9), 899-913. doi:10.1016/s0005-7967(98)00070-9

Wilbur, T. G., & Roscigno, V. J. (2016). First-generation Disadvantage and College Enrollment/Completion. *Socius: Sociological Research for a Dynamic World, 2*(1-11).

Wine, J. (1971). TEST ANXIETY AND DIRECTION OF ATTENTION. *Psychological Bulletin, 76*(2), 92-104. doi:10.1037/h0031332

Wine, J. (1980). Cognitive-Attentional Theory of Test Anxiety. In I. G. Sarason (Ed.), *Test Anxiety: Theory, Research and Applications* (pp. 349-385). Hillsdale, NJ: Erlbaum.

Wise, E. H., & Haynes, S. N. (1983). COGNITIVE TREATMENT OF TEST ANXIETY - RATIONAL RESTRUCTURING VERSUS ATTENTIONAL TRAINING. *Cognitive Therapy and Research, 7*(1), 69-77. doi:10.1007/bf01173425

Wittmaier, B. C. (1972). Test anxiety and study habits. *The Journal of Educational Research, 65*(8), 352-354. doi:10.1080/00220671.1972.10884344

Wolpe, J. (1958). *Psychotherapy by reciprocal inhibition*. Stanford, CA: Stanford University Press.

Wong, S. S. (2008). The Relations of Cognitive Triad, Dysfunctional Attitudes, Automatic Thoughts, and Irrational Beliefs with Test Anxiety. *Current Psychology*, *27*(3), 177-191. doi:10.1007/s12144-008-9033-y

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. doi:10.1080/10615800412331292606

Xu, F. F., Cai, Y., & Tu, D. B. (2021). Psychometric properties of TAS, TAI, FAT test anxiety scales 6 in Chinese university students: a Bifactor IRT study. *Current Psychology*. doi:10.1007/s12144-020-00610-w

Xu, X. B., Xia, M. Y., & Pang, W. G. (2021). Family socioeconomic status and Chinese high school students' test anxiety: Serial mediating role of parental psychological control, learning resources, and student academic self-efficacy. *Scandinavian Journal of Psychology*. doi:10.1111/sjop.12750

Yerdelen, S., McCaffrey, A., & Klassen, R. M. (2016). Longitudinal examination of procrastination and anxiety, and their relation to self-efficacy for self-regulated learning: Latent Growth Curve Modeling. *Kuram ve Uygulamada Eğitim Bilimleri/Educational Sciences: Theory & Practice*, *16*(1), 5-22.

Yilmaz, A. E., Gencoz, T., & Wells, A. (2008). Psychometric Characteristics of the Penn State Worry Questionnaire and Metacognitions Questionnaire-30 and Metacognitive Predictors of Worry and Obsessive-Compulsive

Symptoms in a Turkish Sample. *Clinical Psychology & Psychotherapy, 15*(6), 424-439. doi:10.1002/cpp.589

Zatz, S., & Chassin, L. (1985). Cognitions of test-anxious children under naturalistic test-taking conditions. *Journal of Consulting and Clinical Psychology, 53*(3), 393-401. doi:10.1037/0022-006X.53.3.393

Zeidner, M. (1996). How do high school and college students cope with test situations? *British Journal of Educational Psychology, 66*(1), 115-128. doi:10.1111/j.2044-8279.1996.tb01181.x

Zeidner, M. (1998). *Test anxiety: The state of the art*. New York: Kluwer Academic Publishers.

Zeidner, M., & Matthews, G. (2005). Evaluation anxiety. In A. J. Elliot & C. S. Dweck (Eds.), *Handbook of competence and motivation*. London: Guildford Press.

Zeidner, M., & Matthews, G. (2011). *Anxiety 101*.

Zeidner, M., & Nevo, B. (1993). Test Anxiety Inventory/Hebrew Adaptation (TAI/HB): Scale development, psychometric properties, and some demographic and cognitive correlates. *Megamot, 35*, 293-306.

Zohar, D. (1998). An additive model of test anxiety: Role of exam-specific expectations. *Journal of Educational Psychology, 90*(2), 330-340. doi:10.1037/0022-0663.90.2.330

Appendices

List of Appendices

Appendix A

Search protocols for electronic databases

Overview of eligibility criteria:

- **P**articipants: test anxious undergraduate university students
- **I**nterventions: psychological, educational, or pharmacological interventions
- **C**omparators: waitlist, pill or psychological placebo, TAU, or other active psychological, educational, or pharmacological conditions
- **O**utcomes: change in self-rated test anxiety scores (primary) and change in 'real world' examination scores (secondary)
- **S**tudy design: randomized clinical trials (RCT)

Additionally, eligible studies must be:

- Published in a peer-reviewed journal
- Published in English

No restrictions on date will be made.

Databases to be searched: Scopus, Web of knowledge (Science and Social Sciences Index), Cochrane Central Register of Controlled Trials (CENTRAL), MEDLINE, PsycINFO, Educational Resources Information Center (ERIC).

Additionally, manual searches will be conducted, including reference lists of past reviews and relevant literature (e.g., published RCT articles, books or book chapters on the efficacy of test anxiety interventions). Emails will also be sent to the authors of the previous meta-analytic reviews asking for the list of studies they included in their analyses.

Search filter for: **SCOPUS**

Conducted on: _____

Step	Search Terms	No. Records
1	"Test anxiety" OR "Examination anxiety" OR "Test stress" OR "Examination stress" OR "Test performance" OR "Examination performance" OR "Evaluation Stress" OR "Evaluative Stress" OR "Evaluative Anxiety" OR "Evaluation Anxiety" OR "Test fear" OR "Examination fear"	
2	College student OR University student OR undergraduate	
3	Randomi?ed controlled trial OR Controlled clinical trial OR Clinical trial OR PHASE* Trial OR RCT OR randomi?ed OR random allocation OR randomly allocated OR Controlled trial OR placebo OR outcome study	
4	Psycho* or behavi* or cogniti* or metacogniti* or implosive or desensiti* or interpersonal or gestalt or attenti* or activation* or rational or bibliotherapy or counsel* or supportive or study skill* or skill focus* or testwise* or relax* or training* or nondirective or guided or imag* or computer* or CBT or pharmaco* or drug* or treatment or intervention	
5	#1 AND #2 AND #3 AND #4	
6	Limit #5: English Language	

Notes: Search is to be limited to the Subject Areas: 'Life Sciences, Health Sciences, Social Sciences and Humanities' subject areas, with all document types permissible. Each set of terms to be searched under the 'Article, Title, Abstract, Keywords' heading.

Search filter for: **WEB OF KNOWLEDGE (SCIENCE AND SOCIAL SCIENCE INDEX)**

Conducted on: _____

Step	Search Terms	No. Records
1	"Test anxiety" OR "Examination anxiety" OR "Test stress" OR "Examination stress" OR "Test performance" OR "Examination performance" OR "Evaluation Stress" OR "Evaluative Stress" OR "Evaluative Anxiety" OR "Evaluation Anxiety" OR "Test fear" OR "Examination fear"	
2	College student OR University student OR undergraduate	
3	Randomi?ed controlled trial OR Controlled clinical trial OR Clinical trial OR PHASE* Trial OR RCT OR randomi?ed OR random allocation OR randomly allocated OR Controlled trial OR placebo OR outcome study	
4	Psycho* or behavi* or cogniti* or metacogniti* or implosive or desensiti* or interpersonal or gestalt or attenti* or activation* or rational or bibliotherapy* or counsel* or supportive or study skill* or skill focus* or testwise* or relax* or training* or nondirective or guided or imag* or computer* or CBT or pharmaco* or drug* or treatment or intervention or study	
5	#1 AND #2 AND #3 AND #4	
6	Limit #5: English Language	

Note: Each set of terms is searched under the 'Topic' heading.

Search filter for: **Cochrane Central Register of Controlled Trials (CENTRAL)**

Conducted on: _____

Step	Search Terms	No. Records
1	"Test anxiety" OR "Examination anxiety" OR "Test stress" OR "Examination stress" OR "Test performance" OR "Examination performance" OR "Evaluation Stress" OR "Evaluative Stress" OR "Evaluative Anxiety" OR "Evaluation Anxiety" OR "Test fear" OR "Examination fear"	
2	College student OR University student OR undergraduate	
3	#1 AND #2	
4	Limit #10: Trials	

Notes: (i) no need to use terms for trials as CENTRAL is a repository for them; (ii) no English language limit; (iii) there is a need to limit by 'Trials' heading as CENTRAL will also return protocols and reviews etc. Each set of terms to be searched under the "Title, Abstract, Keywords" heading.

Search filter for: Ovid **MEDLINE**

Conducted on: _____

Step	**Search Terms**	**No. Records**
1	"Test anxiety" OR "Examination anxiety" OR "Test stress" OR "Examination stress" OR "Test performance" OR "Examination performance" OR "Evaluation Stress" OR "Evaluative Stress" OR "Evaluative Anxiety" OR "Evaluation Anxiety" OR "Test fear" OR "Examination fear"	
2	College student OR University student OR undergraduate	
3	Randomized controlled trial OR Randomised controlled trial OR Controlled clinical trial OR Clinical trial OR PHASE* Trial OR RCT OR randomized OR randomised OR random allocation OR randomly allocated OR Controlled trial OR placebo OR outcome study	
4	Psycho* or behavi* or cogniti* or metacogniti* or implosive or desensiti* or interpersonal or gestalt or attenti* or activation* or rational or bibliotherapy* or counsel* or supportive or study skill* or skill focus* or testwise* or relax* or training* or nondirective or guided or imag* or computer* or CBT or pharmaco* or drug* or treatment or intervention	
5	#1 AND #2 AND #3 AND #4	
6	Limit #8 to English	

Notes: Resources to be searched to include Ovid MEDLINE(R) 1946 To March Week 5 2016; Ovid MEDLINE(R) 1996 to March Week 5 2016; Ovid MEDLINE(R) In-Process

& Other Non-Indexed Citations April 14, 2016; Ovid MEDLINE(R) Daily Update April 14, 2016. All document types are permissible. Each search step limited to English language.

Search filter for: **PsycINFO**

Conducted on: _____

Step	Search Terms	No. Records
1	"Test anxiety" OR "Examination anxiety" OR "Test stress" OR "Examination stress" OR "Test performance" OR "Examination performance" OR "Evaluation Stress" OR "Evaluative Stress" OR "Evaluative Anxiety" OR "Evaluation Anxiety" OR "Test fear" OR "Examination fear"	
2	College students OR university students OR undergraduates	
3	Randomi?ed controlled trial OR Controlled clinical trial OR Clinical trial OR PHASE* Trial OR RCT OR randomi?ed OR random allocation OR randomly allocated OR Controlled trial OR placebo OR outcome study	
4	Psycho* or behavi* or cogniti* or metacogniti* or implosive or desensiti* or interpersonal or gestalt or attenti* or activation* or rational or bibliotherapy* or counsel* or supportive or study skill* or skill focus* or testwise* or relax* or training* or nondirective or guided or imag* or computer* or CBT or pharmaco* or drug* or treatment or intervention	
8	#1 AND #2 AND #3 AND #4	
9	Limit #6: English Language	

Search filter for: **Educational Resources Information Center**

Web: https://eric.ed.gov/

Conducted on: _____

Step	Search Terms	No. Records
1	"Test anxiety" (Limit: Peer reviewed only)	
2	Limit #1: Descriptor = "Higher Education"	

Notes: No advanced search options. Terms to be used adopted to be relatively 'inclusive' approach, so that as many potentially relevant records as possible can be included.

Appendix B

Full-text screening assessments

Appendix B. Full-text screening of articles for potential inclusion into the review

#	Study (first author, year)	Peer-reviewed journal?	English language?	Test anxious university students?	Psychological, educational, or pharmacological intervention(s)?	Control condition(s)?	Test anxiety measure?	Randomization?
1.	Akbari 2001	-	N	-	-	-	-	-
2.	Allen 1971; Allen 1974	Y	Y	Y	Y	Y	Y	Y
3.	Allen 1973	Y	Y	Y	Y	Y	Y	Y
4.	Alsina-Jurnet 2007	Y	Y	Y	Y	Y	Y	N
5.	Altmaier 1981	Y	Y	Y	Y	Y	Y	Y
6.	Annis 1986	Y	Y	Y	Y	Y	Y	N
7.	Anton 1976	Y	Y	Y	Y	Y	Y	Y
8.	Aponte 1971	Y	Y	Y	Y	Y	Y	Y
9.	Armstrong 2001	Y	Y	Y	N	Y	N	Y
10.	Arnkoff 1986	Y	Y	Y	Y	Y	Y	Y
11.	Baghurst 2014	Y	Y	Y	Y	Y	Y	N
12.	Bajtelsmit 1977	Y	Y	N	Y	Y	Y	Y
13.	Bajtelsmit 1976	Y	Y	Y	Y	Y	Y	N
14.	Baker 2003	Y	Y	Y	N	Y	Y	Y
15.	Barabasz 1981	Y	Y	Y	Y	Y	Y	Y
16.	Bauman 1994	Y	Y	Y	Y	Y	Y	Y
17.	Bedell 1976	Y	Y	Y	Y	Y	Y	Y
18.	Benor 2009	Y	Y	Y	Y	Y	Y	N

19.	Blank-Spadoni 2014	N	-	-	-	-	-	-
20.	Breedlove 2004	Y	Y	Y	N	Y	N	Y
21.	Brown 2011	Y	Y	Y	Y	N	Y	Y
22.	Bruch 1978	Y	Y	N	Y	Y	Y	Y
23.	Buglione 1990	Y	Y	Y	Y	N	Y	N
24.	Calef 1974	Y	Y	Y	Y	Y	Y	Y
25.	Can 2012	-	N	-	-	-	-	-
26.	Chang-Liang 1976	Y	Y	Y	Y	Y	Y	Y
27.	Chang 2008	Y	Y	N	N	Y	N	Y
28.	Clark 1998	Y	Y	Y	N	Y	Y	Y
29.	Cohen 1969	Y	Y	Y	Y	Y	Y	Y
30.	Cooley 1980	Y	Y	Y	Y	Y	y	Y
31.	Cornish 1973	Y	Y	Y	Y	Y	Y	Y
32.	Counts 1978	Y	Y	Y	Y	Y	Y	Y
33.	Crighton 1969	Y	Y	Y	Y	N	Y	Y
34.	Crowley 1986	Y	Y	Y	Y	Y	Y	Y
35.	D'Alelio 1981	Y	Y	Y	Y	Y	Y	Y
36.	Damer 2011	Y	Y	Y	Y	Y	Y	N
37.	Dawley 1973; Dawley 1973	Y	Y	Y	Y	Y	Y	Y
38.	Dawley 1973	Y	Y	Y	Y	Y	Y	Y
39.	Decker 1987	Y	Y	N	Y	Y	Y	N
40.	Decker 1981	Y	Y	Y	Y	Y	Y	Y
41.	Deffenbacher 1976	Y	Y	Y	Y	Y	Y	N
42.	Deffenbacher 1979; Deffenbacher 1980	Y	Y	Y	Y	Y	Y	Y

43.	Deffenbacher 1980; Deffenbacher 1981	Y	Y	Y	Y	Y	Y	Y
44.	Deffenbacher 1980; Deffenbacher 1981	Y	Y	Y	Y	Y	Y	Y
45.	Deffenbacher 1979	Y	Y	Y	Y	Y	Y	Y
46.	Deffenbacher 1978	Y	Y	Y	Y	N	Y	Y
47.	Deffenbacher 1976	Y	Y	Y	Y	Y	Y	N
48.	Dendato 1986	Y	Y	Y	Y	Y	N	Y
49.	Denney 1977	Y	Y	Y	Y	Y	Y	Y
50.	Denney 1974	Y	Y	Y	Y	Y	Y	Y
51.	DeVaney 2010	Y	Y	N	N	N	N	N
52.	Doctor 1970	Y	Y	N	Y	Y	Y	Y
53.	Donner 1969	Y	Y	Y	Y	Y	Y	Y
54.	Dundas 2013	Y	Y	Y	Y	N	Y	N
55.	Dundas 2009	Y	Y	Y	Y	N	N	N
56.	Emery 1967	Y	Y	Y	Y	Y	N	Y
57.	Fallon 2011	Y	Y	Y	N	Y	N	Y
58.	Finger 1977	Y	Y	Y	Y	Y	Y	Y
59.	Freeling 1970	Y	Y	Y	Y	N	Y	Y
60.	Fulton 1989	Y	Y	N	N	Y	N	Y
61.	Furlan 2013	-	N	-	-	-	-	-
62.	Gallagher 1978	Y	Y	Y	Y	Y	Y	Y
63.	Garlington 1968	Y	Y	Y	Y	Y	Y	Y
64.	Goldburgh 1968	Y	Y	Y	Y	Y	N	N
65.	Goldfried 1978	Y	Y	Y	Y	Y	Y	Y
66.	Gosselin 1995	Y	Y	Y	Y	Y	Y	Y
67.	Guidry 1974	Y	Y	Y	Y	Y	Y	Y

68.	Hahm 2016	Y	Y	Y	Y	Y	Y	Y
69.	Hall 1972	Y	Y	Y	Y	N	Y	Y
70.	Hampel 1997	N	-	-	-	-	-	-
71.	Handelzalts 2010	Y	Y	Y	Y	Y	Y	N
72.	Harris 1980	Y	Y	Y	Y	Y	Y	Y
73.	Harris 1983	Y	Y	Y	Y	Y	Y	Y
74.	Head 1984	Y	Y	N	N	Y	Y	Y
75.	Holahan 1979	Y	Y	Y	Y	Y	Y	Y
76.	Holroyd 1976	Y	Y	Y	Y	Y	Y	Y
77.	Horne 1974	Y	Y	Y	Y	Y	Y	Y
78.	Horne 1977	Y	Y	Y	Y	Y	Y	Y
79.	Hudesman 1984	Y	Y	Y	Y	Y	Y	N
80.	Hussian 1978	Y	Y	Y	Y	Y	Y	Y
81.	Hymen 1978	Y	Y	Y	Y	N	Y	Y
82.	Ihli 1969	Y	Y	Y	Y	N	Y	Y
83.	Jaffe 1972	Y	Y	Y	Y	Y	Y	Y
84.	James 1973	Y	Y	Y	Y	Y	N	Y
85.	Johnson 1978	N	-	-	-	-	-	-
86.	Johnson 1984	Y	Y	N	N	Y	N	Y
87.	Johnson 1968	Y	Y	Y	Y	Y	Y	Y
88.	Kaplan 1979	Y	Y	Y	Y	Y	N	Y
89.	Katahn 1966	Y	Y	Y	Y	Y	Y	N
90.	Katz 1978	N	-	-	-	-	-	-
91.	Kavurmaci 2015	Y	Y	Y	N	Y	Y	N
92.	Kim 2005	N	-	-	-	-	-	-
93.	Kirkland 1980	Y	Y	Y	Y	Y	Y	Y
94.	Kooken 1984	Y	Y	N	Y	Y	Y	Y

95.	Kostka 1974	Y	Y	Y	Y	Y	Y	Y
96.	Kumar 2015	N	Y	N	Y	Y	N	N
97.	La Femina 2000	N	-	-	-	-	-	-
98.	Lent 1978	Y	Y	Y	Y	Y	Y	Y
99.	Levine 1980	Y	Y	Y	Y	Y	Y	Y
100.	Lomont 1971	Y	Y	Y	Y	Y	Y	Y
101.	Lurie 1980	Y	Y	Y	Y	Y	Y	Y
102.	Marchetti 1977	Y	Y	Y	Y	Y	Y	Y
103.	Martin 1999	N	-	-	-	-	-	-
104.	Mason 1983	N	-	-	-	-	-	-
105.	Maxfield 2000	Y	Y	Y	Y	Y	Y	Y
106.	McCordick 1979	Y	Y	Y	Y	Y	Y	Y
107.	McCordick 1981	Y	Y	Y	Y	Y	Y	Y
108.	McGlynn 1978	Y	Y	Y	Y	Y	Y	Y
109.	McManus 1971	Y	Y	Y	Y	Y	N	N
110.	McMillan 1972	Y	Y	Y	Y	Y	Y	N
111.	Meichenbaum 1972	Y	Y	Y	Y	Y	Y	Y
112.	Melnick 1976	Y	Y	Y	Y	Y	Y	Y
113.	Messineo 2015	Y	Y	N	Y	Y	Y	Y
114.	Mitchell 1975	Y	Y	Y	Y	Y	Y	Y
115.	Mitchell 1970	Y	Y	Y	Y	Y	Y	Y
116.	Mitchell 1972	Y	Y	Y	Y	Y	Y	Y
117.	Mohler 2014	N	-	-	-	-	-	-
118.	Muis 2013	Y	Y	N	Y	Y	Y	Y
119.	Myint 2011	Y	Y	N	Y	Y	Y	N
120.	Nadinloyi 2013	N	-	-	-	-	-	-
121.	Naveh-Benjamin 1991	Y	Y	Y	Y	Y	N	Y

122.	Neuderth 2009	Y	Y	Y	N	Y	N	N
123.	Onem 2013	Y	Y	N	N	Y	Y	N
124.	Orbach 2007	Y	Y	Y	Y	Y	Y	Y
125.	Osterhouse 1972	Y	Y	Y	Y	Y	Y	N
126.	Pham 1999	Y	Y	N	N	Y	N	Y
127.	Pintado 2016	Y	Y	Y	Y	Y	Y	N
128.	Prezas 1995	N	-	-	-	-	-	-
129.	Prochaska 1971	Y	Y	Y	Y	Y	Y	Y
130.	Rajiah 2014	Y	Y	Y	Y	Y	Y	Y
131.	Ramirez 2011	Y	Y	N	Y	Y	Y	Y
132.	Reed 1980	Y	Y	Y	Y	Y	Y	Y
133.	Reed 1974	Y	Y	Y	Y	N	Y	Y
134.	Register 1991	Y	Y	Y	Y	Y	Y	Y
135.	Richardson 1974	Y	Y	Y	Y	Y	Y	Y
136.	Ricketts 1984	Y	Y	Y	Y	Y	Y	Y
137.	Rodger 2007	Y	Y	N	N	N	Y	Y
138.	Romano 1978	Y	Y	Y	Y	Y	Y	Y
139.	Russell 1982	Y	Y	Y	Y	Y	Y	Y
140.	Russell 1974	Y	Y	Y	Y	Y	Y	N
141.	Russell 1975	Y	Y	Y	Y	Y	Y	Y
142.	Russell 1973; Russell 1974	Y	Y	Y	Y	Y	Y	N
143.	Russell 1976	Y	Y	Y	Y	Y	Y	Y
144.	Ruthig 2004	Y	Y	N	N	Y	Y	N
145.	Sapp 1996	Y	Y	Y	Y	Y	Y	Y
146.	Saravanan 2014	Y	Y	Y	Y	Y	Y	Y
147.	Sharif 2004	Y	Y	N	Y	Y	N	N

148.	Shu 2011	-	N	-	-	-	-	-
149.	Smith 1973	Y	Y	Y	Y	Y	Y	Y
150.	Snyder 1977	Y	Y	Y	Y	Y	Y	Y
151.	Spies 1979	Y	Y	Y	Y	Y	Y	Y
152.	Suinn 1968	Y	Y	Y	Y	Y	Y	N
153.	Suinn 1973	Y	Y	Y	Y	Y	Y	Y
154.	Tatum 2006	Y	Y	N	Y	Y	N	N
155.	Taylor 1971	Y	Y	Y	Y	Y	Y	Y
156.	Thyer 1981; Himle 1984	Y	Y	Y	Y	N	Y	Y
157.	Topp 1989	Y	Y	Y	Y	Y	Y	N
158.	Turan 2010	-	N	-	-	-	-	-
159.	von der Embse 2015	Y	Y	N	N	N	Y	Y
160.	Wachelka 1999	Y	Y	N	Y	Y	Y	Y
161.	Weiner 1975	Y	Y	N	N	Y	Y	Y
162.	Wise 1983	Y	Y	Y	Y	Y	Y	Y
163.	Wise 1986	Y	Y	N	N	Y	N	Y
164.	Yehuda 2005	Y	Y	Y	N	Y	N	Y
165.	Zimmer 1994	Y	Y	N	Y	Y	Y	Y
166.	Kotter 2016	Y	Y	N	N	Y	Y	Y
167.	Miller 2016	Y	Y	N	Y	N	Y	N
168.	Reiss 2017	Y	Y	Y	Y	Y	Y	Y

Notes.

Y = Yes (meets criterion); N = No (does not meet criterion).

The full completed assessments for each of the above can be found in the supplementary files attached to the published review:

https://www.sciencedirect.com/science/article/abs/pii/S0887618518302032.

Appendix C

Risk of bias assessments

Appendix C: Summary of Risk of Bias assessments.

Study ID	Study (first author, year)	Random sequence generation (selection bias)	Allocation concealment (selection bias)	Blinding of participants and personnel (performance bias)	Blinding of outcome assessment (detection bias)	Incomplete outcome (attrition bias)	Selective reporting (reporting bias)
1.	Mitchell 1970	?	?	+	?	-	?
2.	Lomont 1971	?	?	+	?	-	?
3.	Prochaska 1971	?	?	+	?	-	?
4.	Meichenbaum 1972	?	?	+	?	-	?
5.	Dawley 1973	?	?	+	?	-	?
6.	Kostka 1974	+	?	+	?	+	?
7.	Guidry 1975	?	?	+	?	-	?
8.	Mitchell 1975	?	?	+	?	-	?
9.	Anton 1976	?	?	+	?	-	?
10.	Bedell 1976	?	?	+	?	?	?
11.	Chang-Liang 1976	?	?	+	?	-	?
12.	Melnick 1976	+	?	+	?	+	?
13.	Russell 1976	?	?	+	?	+	?
14.	Denney 1977	?	?	+	-	+	?
15.	Finger 1977	?	?	+	?	-	?
16.	Horne 1977	?	?	+	?	?	?
17.	Counts 1978	?	?	+	?	+	?
18.	Gallagher 1978	?	?	+	?	+	?
19.	Lent 1978	+	?	+	?	+	?

20.	Romano 1978	?	?	+	?	-	?
21.	Deffenbacher 1979a	?	?	+	-	-	?
22.	Deffenbacher 1979b	?	?	+	?	-	?
23.	Holahan 1979	?	?	+	?	-	?
24.	Deffenbacher 1980a	+	?	+	?	-	?
25.	Deffenbacher 1980b	?	?	+	-	-	?
26.	Levine 1980	?	?	+	?	-	?
27.	Lurie 1980	?	?	+	?	+	?
28.	Reed 1980	?	?	+	?	-	?
29.	Altmaier 1981	?	?	+	?	?	?
30.	Barabasz 1981	?	?	+	?	-	?
31.	D'Alelio 1981	?	?	+	?	-	?
32.	Decker 1981	?	?	+	?	-	?
33.	McCordick 1981	?	?	+	?	-	?
34.	Russell 1982	?	?	+	?	+	?
35.	Ricketts 1984	?	?	+	?	?	?
36.	Crowley 1986	?	?	+	?	+	?
37.	Bauman 1994	?	?	+	?	?	?
38.	Sapp 1996	?	?	+	-	-	?
39.	Maxfield 2000	?	?	+	?	-	?
40.	Orbach 2007	-	?	+	?	+	?
41.	Raijah 2014	-	?	+	?	-	?
42.	Saravanan 2014	?	-	+	?	-	?

43.	Hahm 2016	?	?	+	-	?	?
44.	Reiss 2017	?	?	+	?	+	?

Notes.

+ = high risk of bias; - = low risk of bias; ? = unknown risk of bias.

Appendix D

Supplementary forest and funnel plots

Appendix D: Annotated forest and funnel plots.

The following forest and funnel plots are contained with this document (all produced using Revman):

Figure D.1. Forest plot of intervention vs. control for reducing test anxiety severity (at post-treatment)

Figure D.2. Funnel plot of intervention vs. control for reducing test anxiety severity (at post-treatment)

Figure D.3. Forest plot of intervention vs. control for reducing test anxiety severity (at follow-up)

Figure D.4. Funnel plot of intervention vs. control for reducing test anxiety severity (at follow-up)

Figure D.5. Forest plot of intervention vs. control for reducing in-situ (state) test anxiety severity

Figure D.6. Funnel plot of intervention vs. control for reducing in-situ (state) test anxiety severity

Figure D.7. Forest plot of intervention vs. control for improving academic performance (at post-treatment)

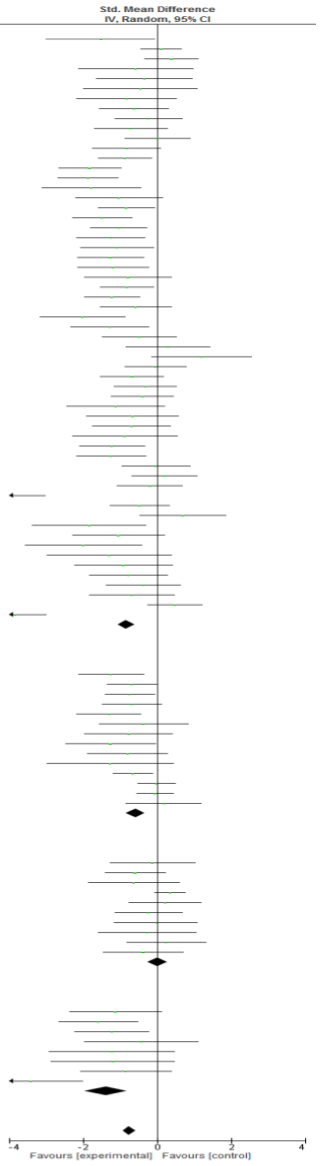
Figure D.8. Funnel plot of intervention vs. control for improving academic performance (at post-treatment)

Figure D.9. Forest plot of intervention vs. control for improving academic performance (at follow-up)

Figure D.10. Funnel plot of intervention vs. control for improving academic performance (at follow-up)

Figure D.1. Forest plot of intervention vs. control for reducing test anxiety severity (at post-treatment)

Study or Subgroup	Experimental Mean	Experimental SD	Experimental Total	Control Mean	Control SD	Control Total	Weight	Std. Mean Difference IV, Random, 95% CI
1.1.1 BT								
Altmaier 1981	41.1	7.92	10	54	7.73	3	0.8%	-1.52 [-2.99, -0.05]
Anton 1976	45.5	8.84	32	44.36	11.26	22	1.5%	0.11 [-0.43, 0.66]
Bauman 1994	51.33	13.23	15	47	7.41	15	1.4%	0.39 [-0.33, 1.12]
Bedell, 1976 (1)	22.9	5.04	10	26.1	6.33	2	0.7%	-0.57 [-2.11, 0.97]
Bedell, 1976 (2)	23.5	7.15	10	26.1	6.33	3	0.9%	-0.35 [-1.65, 0.95]
Bedell, 1976 (3)	21.9	8.84	10	26.1	6.33	2	0.7%	-0.45 [-1.98, 1.08]
Bedell, 1976 (4)	20.7	5.98	10	26.1	6.33	3	0.9%	-0.83 [-2.18, 0.51]
Chang-Liang 1976 (5)	68.45	27.58	20	86.78	31.94	6	1.2%	-0.62 [-1.55, 0.31]
Chang-Liang 1976 (6)	79.91	28.98	21	86.78	31.94	6	1.2%	-0.23 [-1.13, 0.68]
Chang-Liang 1976 (7)	66.63	26.64	23	86.78	31.94	5	1.1%	-0.71 [-1.70, 0.28]
Counts 1978 (8)	24.2	6.88	10	24.15	3.74	10	1.2%	0.01 [-0.97, 0.89]
Counts 1978 (9)	18.8	7.98	10	24.15	3.74	10	1.2%	-0.82 [-1.74, 0.10]
Dawley 1973	209.17	29.03	12	238	33.78	24	1.4%	-0.87 [-1.60, -0.15]
Deffenbacher 1979a (10)	22.77	7.41	15	36.08	6.87	17	1.3%	-1.82 [-2.66, -0.98]
Deffenbacher 1979a (11)	23.16	6.68	17	36.08	6.87	18	1.3%	-1.86 [-2.67, -1.05]
Deffenbacher 1979b (12)	16.41	6.7	9	28.2	5.2	5	0.9%	-1.77 [-3.11, -0.43]
Deffenbacher 1979b (13)	20.36	7.92	9	28.2	5.2	5	1.0%	-1.03 [-2.21, 0.15]
Deffenbacher 1980a	24.92	7.89	14	30.72	5.51	15	1.3%	-0.83 [-1.60, -0.07]
Deffenbacher 1980b (14)	28.4	5.49	15	36.85	5.63	16	1.3%	-1.48 [-2.29, -0.67]
Deffenbacher 1980b (15)	29.46	8.03	16	36.85	5.63	15	1.3%	-1.03 [-1.79, -0.28]
Denney 1977 (16)	53.8	29.8	13	91.78	28.41	10	1.2%	-1.25 [-2.17, -0.34]
Denney 1977 (17)	53.8	38.8	9	91.78	28.41	10	1.1%	-1.08 [-2.06, -0.10]
Denney 1977 (18)	57.3	25	14	91.78	28.41	10	1.2%	-1.26 [-2.16, -0.36]
Denney 1977 (19)	51.8	35	11	91.78	28.41	9	1.1%	-1.19 [-2.16, -0.22]
Finger 1977	30.02	6.93	12	35.49	5.4	4	1.0%	-0.79 [-1.95, 0.39]
Gallagher 1978	20.2	7.18	12	25.2	5.28	24	1.4%	-0.92 [-1.54, -0.10]
Guidry 1975	128.08	25.67	12	165.5	32.02	24	1.3%	-1.22 [-1.97, -0.46]
Holahan 1979	18.67	7.76	15	23.08	6.05	6	1.1%	-0.58 [-1.54, 0.39]
Horne 1977 (20)	11.05	5.07	20	21.35	4.28	5	1.0%	-2.02 [-3.17, -0.86]
Horne 1977 (21)	17.05	3.01	20	21.35	4.28	5	1.1%	-1.27 [-2.33, -0.22]
Horne 1977 (22)	18.9	5.07	20	21.35	4.28	5	1.1%	-0.48 [-1.47, 0.51]
Kostka 1974 (23)	35	17.89	6	28.85	20.42	6	1.0%	0.30 [-0.85, 1.44]
Kostka 1974 (24)	51.6	12.18	5	28.85	20.42	6	0.9%	1.21 [-0.14, 2.55]
Levine 1980 (25)	5.07	1.12	15	5.12	1.36	9	1.3%	-0.04 [-0.87, 0.79]
Levine 1980 (26)	4.15	1.41	13	5.12	1.36	10	1.2%	-0.67 [-1.52, 0.18]
Levine 1980 (27)	4.78	0.8	14	5.12	1.36	9	1.3%	-0.31 [-1.16, 0.53]
Lomont 1971	32.5	8.17	8	35.47	6.61	17	1.3%	-0.40 [-1.25, 0.45]
Lurie 1980 (28)	17.33	6.25	6	25.57	7.25	5	0.9%	-1.12 [-2.45, 0.20]
Lurie 1980 (29)	21.5	4.89	8	25.57	7.25	4	0.9%	-0.66 [-1.90, 0.58]
Maxfield 2000	46.71	15.16	7	56	10.14	8	1.1%	-0.69 [-1.74, 0.37]
Meichenbaum 1972	26	4.1	8	30	4.6	3	0.8%	-0.87 [-2.27, 0.53]
Melnick 1976	134.89	28.52	9	163.67	19.97	18	1.2%	-1.21 [-2.08, -0.34]
Mitchell 1970	28.45	3	31	32.42	3.87	6	1.2%	-1.24 [-2.16, -0.31]
Prochaska 1971 (30)	-2	7.72	13	-1.66	9.73	7	1.2%	-0.04 [-0.96, 0.88]
Prochaska 1971 (31)	-0.23	4.19	13	-1.66	9.73	8	1.2%	0.20 [-0.68, 1.09]
Prochaska 1971 (32)	-3.67	10.23	13	-1.66	9.73	8	1.2%	-0.19 [-1.08, 0.69]
Rajjah 2014	20.2	5	20	39.5	4	20	1.0%	-4.18 [-5.32, -3.03]
Reed 1980	28.8	7	9	31.95	5.98	18	1.3%	-0.47 [-1.20, 0.34]
Ricketts 1984	86.54	26.84	13	67.85	20.76	4	1.0%	0.69 [-0.46, 1.84]
Romano 1978 (33)	109.2	30.11	10	167.8	27.6	3	0.7%	-1.84 [-3.38, -0.29]
Romano 1978 (34)	127.6	38.4	10	167.8	27.6	4	0.9%	-1.05 [-2.29, 0.20]
Romano 1978 (35)	116.2	23.33	10	167.8	27.6	3	0.7%	-1.99 [-3.57, -0.41]
Russell 1976 (36)	7	4.9	5	13.7	3.6	3	0.7%	-1.29 [-2.98, 0.39]
Russell 1976 (37)	9.7	4.2	7	13.7	3.6	4	0.9%	-0.91 [-2.23, 0.41]
Russell 1982 (38)	8.2	3.9	10	11.4	4.05	6	1.1%	-0.77 [-1.82, 0.29]
Russell 1982 (39)	9.8	4	9	11.4	4.05	7	1.1%	-0.38 [-1.38, 0.62]
Russell 1982 (40)	7.8	5.8	6	11.4	4.05	7	1.0%	-0.68 [-1.81, 0.45]
Sapp 1996	52.3	8.63	15	48.58	6.4	15	1.4%	0.48 [-0.25, 1.20]
Saravanan 2014	19.28	5.56	32	38.54	4.29	33	1.3%	-3.84 [-4.68, -3.00]
Subtotal (95% CI)			**771**			**565**	**64.9%**	**-0.83 [-1.06, -0.60]**
Heterogeneity: Tau² = 0.55; Chi² = 192.17, df = 58 (P < 0.00001); I² = 70%								
Test for overall effect: Z = 7.00 (P < 0.00001)								
1.1.2 CBT								
Barabasz 1981	3.71	3.2	18	7.81	3.2	9	1.2%	-1.24 [-2.12, -0.37]
Crowley 1986 (41)	45.91	14.02	31	55.5	12.94	13	1.4%	-0.69 [-1.35, -0.02]
Crowley 1986 (42)	47.33	9.84	29	55.5	12.94	14	1.4%	-0.73 [-1.39, -0.08]
D'Alelio 1981 (43)	48	12.5	18	56.6	11.5	10	1.3%	-0.69 [-1.48, 0.11]
D'Alelio 1981 (44)	42.5	9.9	17	56.6	11.5	10	1.2%	-1.30 [-2.17, -0.43]
Decker 1981	21	8.7	9	24.3	8.2	4	1.0%	-0.36 [-1.55, 0.83]
Finger 1977 (45)	28.84	8.85	11	35.49	5.4	4	1.0%	-0.76 [-1.95, 0.42]
Finger 1977 (46)	29.98	3.78	13	35.49	5.4	4	0.9%	-1.26 [-2.48, -0.04]
Holahan 1979	18.33	5.48	12	23.08	6.05	5	1.0%	-0.80 [-1.89, 0.29]
Meichenbaum 1972	24.5	3.8	8	30	4.6	2	0.7%	-1.27 [-2.98, 0.44]
Orbach 2007	47.31	9.49	30	54.25	11.31	28	1.5%	-0.66 [-1.19, -0.13]
Reiss 2017 (47)	56.8	13.08	42	57.03	13.36	24	1.6%	-0.02 [-0.52, 0.48]
Reiss 2017 (48)	56.3	10.52	48	57.03	13.36	24	1.6%	-0.06 [-0.55, 0.43]
Ricketts 1984	73.4	30.81	15	67.85	20.76	5	1.1%	0.18 [-0.83, 1.20]
Subtotal (95% CI)			**301**			**156**	**16.9%**	**-0.58 [-0.83, -0.33]**
Heterogeneity: Tau² = 0.06; Chi² = 18.50, df = 13 (P = 0.14); I² = 30%								
Test for overall effect: Z = 4.47 (P < 0.00001)								
1.1.3 Educational								
Altmaier 1981	52.54	12.85	11	54	7.73	4	1.0%	-0.12 [-1.26, 1.03]
Barabasz 1981	5.93	3	18	7.81	3.2	9	1.3%	-0.59 [-1.41, 0.22]
Decker 1981	18.1	9.5	8	24.3	8.2	4	0.9%	-0.63 [-1.87, 0.61]
Hahm 2016	51.9	9.88	43	48.43	9.9	46	1.6%	0.35 [-0.07, 0.77]
Horne 1977	22.1	3.27	20	21.35	4.28	5	1.1%	0.21 [-0.77, 1.19]
Lent 1978	24.9	5.7	10	26.3	5.6	9	1.2%	-0.24 [-1.14, 0.67]
Lurie 1980	25.37	5.57	8	25.57	7.25	5	1.0%	-0.03 [-1.15, 1.09]
McCordick 1981	-1.11	4.57	9	0.22	4.35	3	0.9%	-0.27 [-1.58, 1.04]
Mitchell 1975	32.8	2.4	20	32.2	1.6	4	1.1%	0.25 [-0.83, 1.33]
Ricketts 1984	60.6	17.91	20	67.85	20.76	4	1.1%	-0.38 [-1.46, 0.70]
Subtotal (95% CI)			**167**			**93**	**11.2%**	**0.02 [-0.25, 0.28]**
Heterogeneity: Tau² = 0.00; Chi² = 6.98, df = 9 (P = 0.64); I² = 0%								
Test for overall effect: Z = 0.12 (P = 0.90)								
1.1.4 Combined								
Altmaier 1981	44.36	8.15	11	54	7.73	4	0.9%	-1.13 [-2.36, 0.11]
Lent 1978 (49)	15.7	6.9	11	26.3	5.6	8	1.1%	-1.58 [-2.65, -0.51]
Lent 1978 (50)	18.4	6.6	10	26.3	5.6	9	1.1%	-1.23 [-2.23, -0.23]
McCordick 1981 (51)	-1.9	4.58	10	0.22	4.35	2	0.7%	-0.43 [-1.96, 1.10]
McCordick 1981 (52)	-5.38	4.14	8	0.22	4.35	2	0.7%	-1.21 [-2.91, 0.48]
McCordick 1981 (53)	-6.78	5.45	9	0.22	4.35	2	0.7%	-1.20 [-2.85, 0.46]
Mitchell 1975 (54)	30.9	1.5	24	32.2	1.6	3	0.9%	-0.84 [-2.06, 0.39]
Mitchell 1975 (55)	26	1.8	40	32.2	1.6	3	0.8%	-3.40 [-4.79, -2.00]
Subtotal (95% CI)			**123**			**33**	**7.0%**	**-1.38 [-1.96, -0.81]**
Heterogeneity: Tau² = 0.24; Chi² = 10.76, df = 7 (P = 0.15); I² = 35%								
Test for overall effect: Z = 4.72 (P < 0.00001)								
Total (95% CI)			**1362**			**847**	**100.0%**	**-0.76 [-0.93, -0.58]**
Heterogeneity: Tau² = 0.43; Chi² = 267.65, df = 90 (P < 0.00001); I² = 66%								
Test for overall effect: Z = 8.56 (P < 0.00001)								
Test for subgroup differences: Chi² = 30.93, df = 3 (P < 0.00001), I² = 90.3%								



Footnotes
(1) RT-Low
(2) RT High
(3) SD Low
(4) SD High
(5) RT
(6) AR
(7) SD
(8) EMG+CCR
(9) CCR
(10) SCR
(11) SD
(12) SCD
(13) SD
(14) AMT
(15) SCD
(16) SD-P
(17) SCD-A
(18) SCD-P
(19) SD-A
(20) MOD
(21) SD
(22) FL
(23) CM
(24) SD
(25) NP
(26) SD
(27) NP+HW
(28) CRI
(29) CRH
(30) IMP-GA
(31) IMP-Dyn
(32) IMP-Sym
(33) SD
(34) EMG-SD
(35) EMG
(36) SD
(37) CCR
(38) SD
(39) SD+CCR
(40) CCR
(41) ICT
(42) gCT
(43) CT-4
(44) CT-8
(45) AT
(46) AT+RT
(47) CBT+IR
(48) CBT
(49) SD+SST
(50) CCD+SST
(51) CBT+SST
(52) CT+SST
(53) SD+SST
(54) SD+SST
(55) RT+SST

Figure D.2. Funnel plot of intervention vs. control for reducing test anxiety severity (at post-treatment)

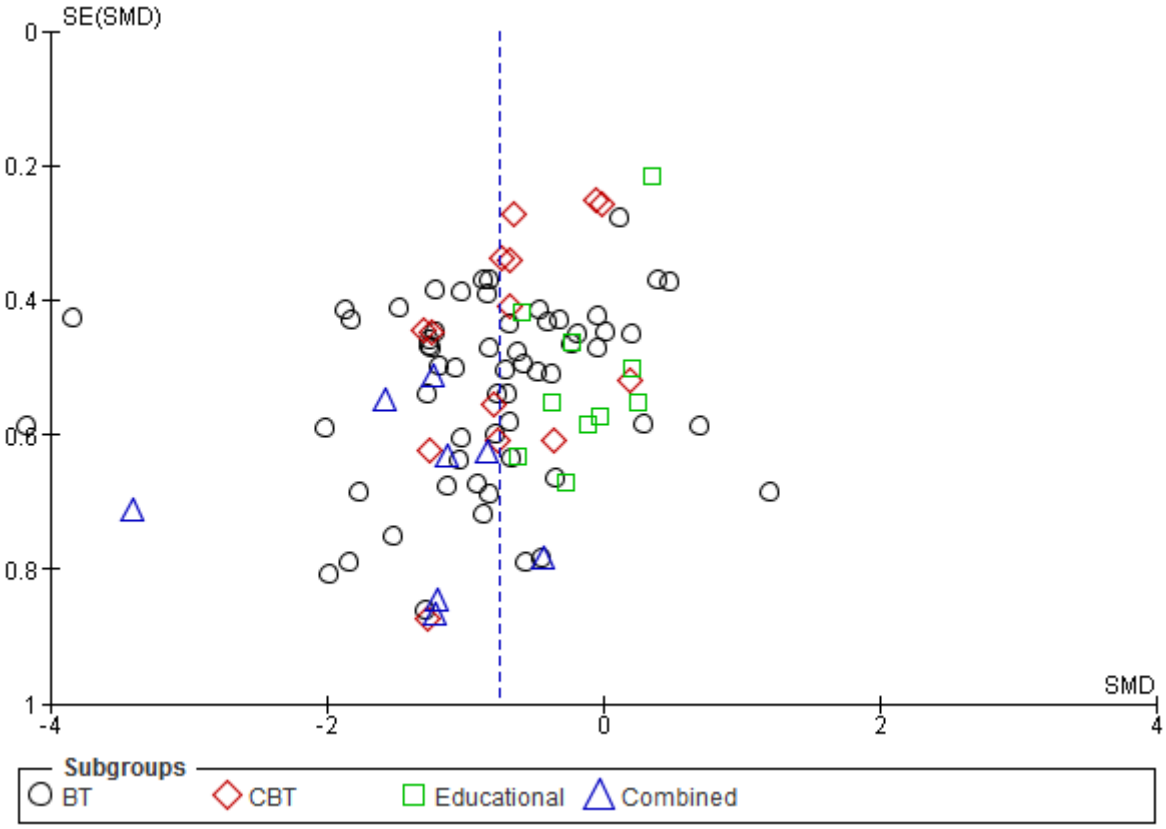
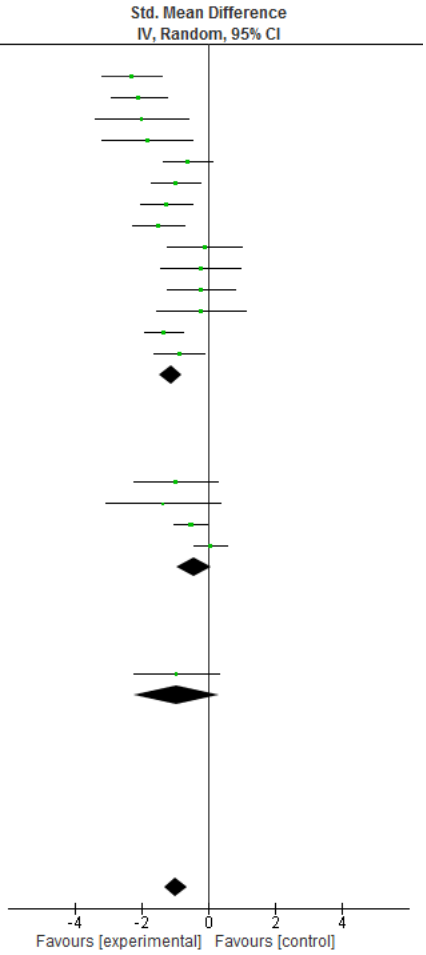


Figure D.3. Forest plot of intervention vs. control for reducing test anxiety severity (at follow-up)

Study or Subgroup	Experimental Mean	Experimental SD	Experimental Total	Control Mean	Control SD	Control Total	Weight	Std. Mean Difference IV, Random, 95% CI
2.1.1 BT								
Deffenbacher 1979a (1)	20.92	7.08	15	37.1	6.75	17	5.4%	-2.28 [-3.20, -1.37]
Deffenbacher 1979a (2)	21.69	7.77	17	37.1	6.75	18	5.8%	-2.07 [-2.91, -1.23]
Deffenbacher 1979b (3)	18.09	5.32	9	28.35	3.58	5	3.5%	-2.00 [-3.40, -0.60]
Deffenbacher 1979b (4)	16.95	6.76	9	28.35	3.58	5	3.7%	-1.81 [-3.16, -0.46]
Deffenbacher 1980a	25.39	6.95	14	29.13	5.15	15	6.3%	-0.60 [-1.34, 0.15]
Deffenbacher 1980b (5)	29.8	8.86	16	37.18	5.45	15	6.3%	-0.97 [-1.72, -0.22]
Deffenbacher 1980b (6)	29.39	6.78	15	37.18	5.45	16	6.1%	-1.24 [-2.02, -0.46]
Guidry 1975	109.75	35.94	12	160.96	32.27	24	6.1%	-1.49 [-2.28, -0.71]
Kostka 1974 (7)	18.47	50.33	6	22.65	22.58	6	4.5%	-0.10 [-1.23, 1.03]
Kostka 1974 (8)	13.55	51.8	5	22.65	22.58	6	4.2%	-0.22 [-1.41, 0.98]
Maxfield 2000	39.71	12.88	7	42.5	12.76	8	5.0%	-0.20 [-1.22, 0.81]
Meichenbaum 1972	27	4.5	8	28	4.9	3	3.7%	-0.20 [-1.53, 1.13]
Mitchell 1970	26.61	4.18	31	31.86	3.65	28	7.3%	-1.32 [-1.88, -0.75]
Sapp 1996	46.83	7.5	15	53.52	7.65	15	6.3%	-0.86 [-1.61, -0.11]
Subtotal (95% CI)			**179**			**181**	**74.2%**	**-1.12 [-1.47, -0.77]**
Heterogeneity: Tau² = 0.23; Chi² = 27.81, df = 13 (P = 0.010); I² = 53%								
Test for overall effect: Z = 6.19 (P < 0.00001)								
2.1.2 CBT								
Decker 1981	16	8	9	24.3	8.2	4	4.0%	-0.96 [-2.22, 0.30]
Meichenbaum 1972	22.5	3.5	8	28	4.9	2	2.7%	-1.34 [-3.06, 0.38]
Reiss 2017	47.96	10.94	48	54.04	13.56	24	7.6%	-0.51 [-1.00, -0.01]
Reiss 2017	55.05	11.45	42	54.04	13.56	24	7.6%	0.08 [-0.42, 0.58]
Subtotal (95% CI)			**107**			**54**	**21.9%**	**-0.40 [-0.92, 0.12]**
Heterogeneity: Tau² = 0.11; Chi² = 5.34, df = 3 (P = 0.15); I² = 44%								
Test for overall effect: Z = 1.52 (P = 0.13)								
2.1.3 Educational								
Decker 1981	14.3	10.3	8	24.3	8.2	4	3.9%	-0.95 [-2.24, 0.34]
Subtotal (95% CI)			**8**			**4**	**3.9%**	**-0.95 [-2.24, 0.34]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.45 (P = 0.15)								
2.1.4 Combined								
Subtotal (95% CI)			**0**			**0**		**Not estimable**
Heterogeneity: Not applicable								
Test for overall effect: Not applicable								
Total (95% CI)			**294**			**239**	**100.0%**	**-0.97 [-1.31, -0.64]**
Heterogeneity: Tau² = 0.32; Chi² = 49.38, df = 18 (P < 0.0001); I² = 64%								
Test for overall effect: Z = 5.69 (P < 0.00001)								
Test for subgroup differences: Chi² = 5.00, df = 2 (P = 0.08), I² = 60.0%								



Footnotes
(1) SD
(2) SCR
(3) SCD
(4) SD
(5) AMT
(6) SCD
(7) CM
(8) SD

Figure D.4. Funnel plot of intervention vs. control for reducing test anxiety severity (at follow-up)

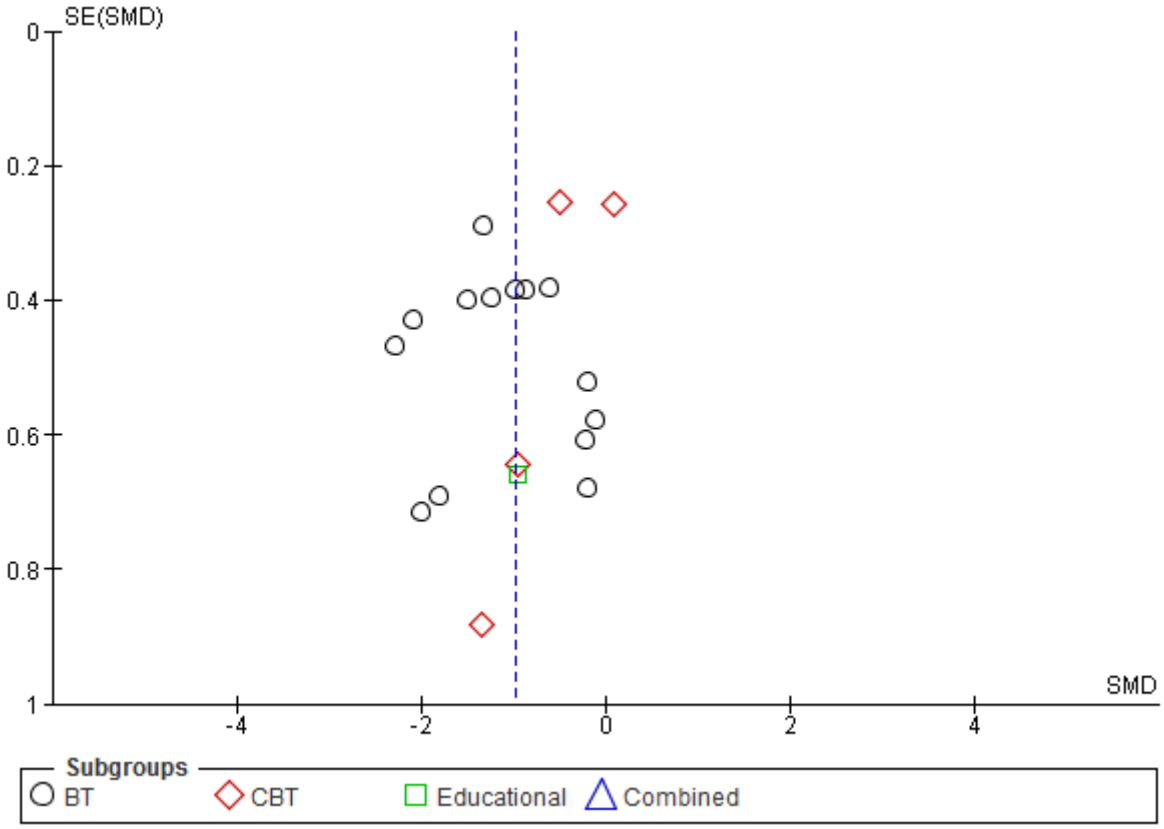
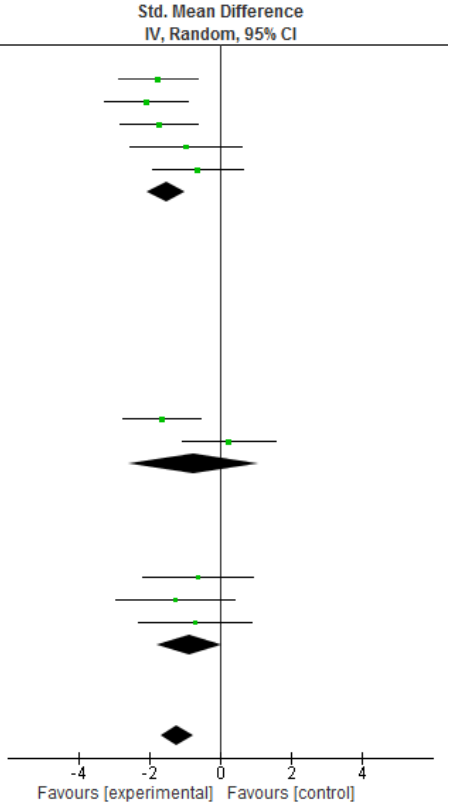


Figure D.5. Forest plot of intervention vs. control for reducing in-situ (state) test anxiety severity

Study or Subgroup	Experimental Mean	Experimental SD	Experimental Total	Control Mean	Control SD	Control Total	Weight	Std. Mean Difference IV, Random, 95% CI
3.1.1 BT								
Horne 1977 (1)	14.35	4.47	20	22.9	5.79	5	12.9%	-1.75 [-2.86, -0.64]
Horne 1977 (2)	15.4	2.77	20	22.9	5.79	5	12.0%	-2.08 [-3.24, -0.92]
Horne 1977 (3)	15.2	3.94	20	22.9	5.79	5	13.0%	-1.72 [-2.83, -0.61]
Russell 1976 (4)	56.6	12.1	5	69	9.6	3	7.3%	-0.95 [-2.53, 0.62]
Russell 1976 (5)	62.3	9.8	7	69	9.6	4	10.5%	-0.63 [-1.90, 0.64]
Subtotal (95% CI)			**72**			**22**	**55.6%**	**-1.52 [-2.06, -0.97]**
Heterogeneity: Tau² = 0.00; Chi² = 3.56, df = 4 (P = 0.47); I² = 0%								
Test for overall effect: Z = 5.46 (P < 0.00001)								
3.1.2 CBT								
Subtotal (95% CI)			**0**			**0**		**Not estimable**
Heterogeneity: Not applicable								
Test for overall effect: Not applicable								
3.1.3 Educational								
Horne 1977	16.5	3.22	20	22.9	5.79	5	13.2%	-1.63 [-2.73, -0.53]
McCordick 1981	6.57	3.1	9	5.75	2.63	3	9.9%	0.25 [-1.06, 1.56]
Subtotal (95% CI)			**29**			**8**	**23.1%**	**-0.73 [-2.57, 1.12]**
Heterogeneity: Tau² = 1.39; Chi² = 4.65, df = 1 (P = 0.03); I² = 79%								
Test for overall effect: Z = 0.77 (P = 0.44)								
3.1.4 Combined								
McCordick 1981	4	2.58	10	5.75	2.63	2	7.5%	-0.62 [-2.17, 0.92]
McCordick 1981	3.43	1.51	9	5.75	2.63	2	6.6%	-1.27 [-2.94, 0.40]
McCordick 1981	3.33	3.2	8	5.75	2.63	2	7.1%	-0.70 [-2.30, 0.90]
Subtotal (95% CI)			**27**			**6**	**21.2%**	**-0.85 [-1.77, 0.08]**
Heterogeneity: Tau² = 0.00; Chi² = 0.36, df = 2 (P = 0.84); I² = 0%								
Test for overall effect: Z = 1.79 (P = 0.07)								
Total (95% CI)			**128**			**36**	**100.0%**	**-1.20 [-1.66, -0.75]**
Heterogeneity: Tau² = 0.10; Chi² = 11.04, df = 9 (P = 0.27); I² = 18%								
Test for overall effect: Z = 5.16 (P < 0.00001)								
Test for subgroup differences: Chi² = 1.89, df = 2 (P = 0.39), I² = 0%								



Footnotes
(1) SD
(2) FL
(3) MOD
(4) SD
(5) CCR

Figure D.6. Funnel plot of intervention vs. control for reducing in-situ (state) test anxiety severity

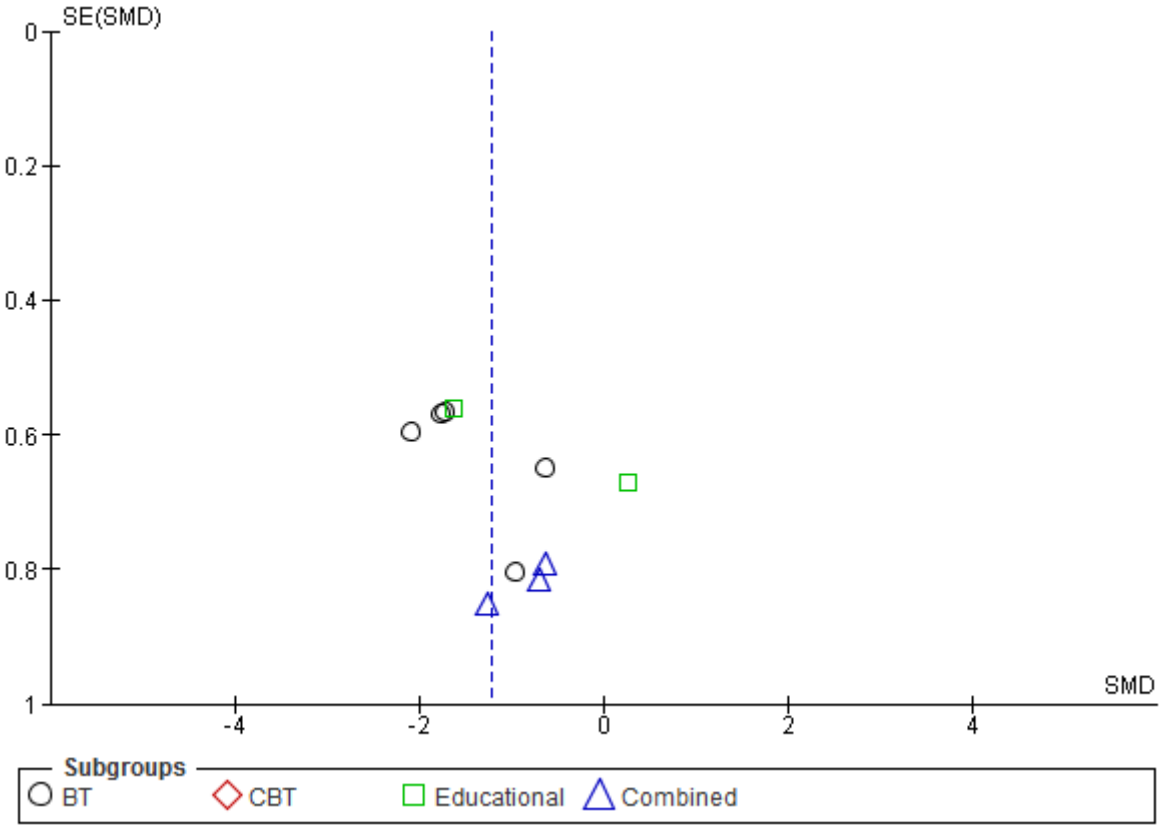
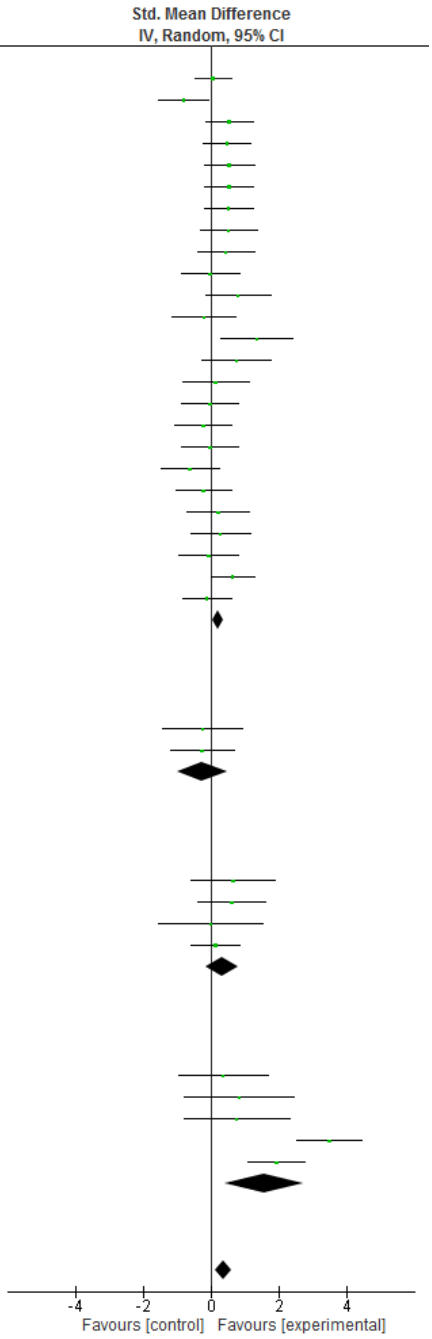


Figure D.7. Forest plot of intervention vs. control for improving academic performance (at post-treatment)

Study or Subgroup	Experimental Mean	Experimental SD	Experimental Total	Control Mean	Control SD	Control Total	Weight	Std. Mean Difference IV, Random, 95% CI
4.1.1 BT								
Anton 1976	2.49	0.77	32	2.43	0.75	22	3.8%	0.08 [-0.47, 0.62]
Bauman 1994	25.7	5.96	15	30.32	5.32	15	3.2%	-0.80 [-1.54, -0.05]
Deffenbacher 1979a (1)	2.76	0.56	17	2.41	0.66	17	3.4%	0.56 [-0.13, 1.25]
Deffenbacher 1979a (2)	2.73	0.59	15	2.41	0.66	18	3.3%	0.50 [-0.20, 1.19]
Deffenbacher 1980a	3.07	0.73	14	2.6	0.91	15	3.2%	0.55 [-0.19, 1.30]
Deffenbacher 1980b	2.5	0.73	16	2.03	0.92	15	3.3%	0.55 [-0.17, 1.27]
Deffenbacher 1980b	2.47	0.64	15	2.03	0.92	16	3.3%	0.54 [-0.18, 1.26]
Denney 1977 (3)	2.97	0.78	14	2.56	0.66	10	3.0%	0.54 [-0.29, 1.37]
Denney 1977 (4)	2.89	0.74	13	2.56	0.66	10	3.0%	0.45 [-0.39, 1.29]
Denney 1977 (5)	2.55	0.74	11	2.56	0.66	10	2.9%	-0.01 [-0.87, 0.84]
Denney 1977 (6)	3.07	0.53	9	2.56	0.66	9	2.6%	0.81 [-0.16, 1.78]
Holahan 1979	2.87	0.75	15	3.01	0.58	6	2.7%	-0.19 [-1.14, 0.76]
Horne 1977 (7)	3.09	0.77	20	2.05	0.58	5	2.4%	1.36 [0.30, 2.42]
Horne 1977 (8)	2.65	0.79	20	2.05	0.58	5	2.5%	0.77 [-0.24, 1.77]
Horne 1977 (9)	2.15	0.65	20	2.05	0.58	5	2.6%	0.15 [-0.83, 1.13]
Levine 1980 (10)	2.89	0.62	13	2.9	0.52	10	3.0%	-0.02 [-0.84, 0.81]
Levine 1980 (11)	2.76	0.72	14	2.9	0.52	9	3.0%	-0.21 [-1.05, 0.63]
Levine 1980 (12)	2.89	0.62	15	2.9	0.52	9	3.0%	-0.02 [-0.84, 0.81]
Lomont 1971	59.13	6.69	8	65.35	10.92	17	2.9%	-0.61 [-1.47, 0.25]
Melnick 1976	69.89	12.55	9	72.56	12.94	18	3.1%	-0.20 [-1.00, 0.60]
Prochaska 1971 (13)	0.35	2.89	13	-0.29	1.87	7	2.7%	0.24 [-0.69, 1.16]
Prochaska 1971 (14)	0.26	1.7	13	-0.29	1.87	8	2.8%	0.30 [-0.59, 1.19]
Prochaska 1971 (15)	-0.45	3.89	13	-0.29	1.87	8	2.8%	-0.05 [-0.93, 0.83]
Raijah 2014	2.9	0.3	20	2.7	0.3	20	3.5%	0.65 [0.02, 1.29]
Sapp 1996	54.19	2.58	15	54.63	4.73	15	3.3%	-0.11 [-0.83, 0.60]
Subtotal (95% CI)			**379**			**299**	**75.3%**	**0.22 [0.05, 0.40]**
Heterogeneity: Tau² = 0.04; Chi² = 29.34, df = 24 (P = 0.21); I² = 18%								
Test for overall effect: Z = 2.46 (P = 0.01)								
4.1.2 CBT								
Decker 1981	1.58	0.78	9	1.79	0.89	4	2.1%	-0.24 [-1.42, 0.94]
Holahan 1979	2.88	0.47	12	3.01	0.58	7	2.7%	-0.24 [-1.18, 0.69]
Subtotal (95% CI)			**21**			**11**	**4.9%**	**-0.24 [-0.98, 0.49]**
Heterogeneity: Tau² = 0.00; Chi² = 0.00, df = 1 (P = 1.00); I² = 0%								
Test for overall effect: Z = 0.65 (P = 0.52)								
4.1.3 Educational								
Decker 1981	2.19	0.31	8	1.79	0.89	4	2.0%	0.67 [-0.58, 1.91]
Horne 1977	2.55	0.8	20	2.05	0.58	5	2.6%	0.63 [-0.37, 1.63]
McCordick 1981	-0.12	1.31	9	-0.13	0.58	2	1.6%	0.01 [-1.52, 1.54]
Mitchell 1975	19.8	18.9	20	16.63	21.79	12	3.3%	0.15 [-0.56, 0.87]
Subtotal (95% CI)			**57**			**23**	**9.4%**	**0.34 [-0.16, 0.84]**
Heterogeneity: Tau² = 0.00; Chi² = 1.03, df = 3 (P = 0.79); I² = 0%								
Test for overall effect: Z = 1.34 (P = 0.18)								
4.1.4 Combined								
McCordick 1981	0.2	0.84	9	-0.13	0.58	3	1.9%	0.38 [-0.94, 1.70]
McCordick 1981	0.71	0.94	8	-0.13	0.58	2	1.5%	0.84 [-0.78, 2.46]
McCordick 1981	0.5	0.76	10	-0.13	0.58	2	1.5%	0.78 [-0.78, 2.35]
Mitchell 1975	56.5	6	40	16.63	21.79	11	2.6%	3.50 [2.53, 4.48]
Mitchell 1975	43.5	6.6	24	16.63	21.79	12	3.0%	1.94 [1.10, 2.78]
Subtotal (95% CI)			**91**			**30**	**10.4%**	**1.58 [0.41, 2.76]**
Heterogeneity: Tau² = 1.38; Chi² = 19.12, df = 4 (P = 0.0007); I² = 79%								
Test for overall effect: Z = 2.64 (P = 0.008)								
Total (95% CI)			**548**			**363**	**100.0%**	**0.37 [0.14, 0.61]**
Heterogeneity: Tau² = 0.30; Chi² = 90.17, df = 35 (P < 0.00001); I² = 61%								
Test for overall effect: Z = 3.12 (P = 0.002)								
Test for subgroup differences: Chi² = 6.86, df = 3 (P = 0.08), I² = 56.3%								



Footnotes
(1) SCR
(2) SD
(3) SCP-P
(4) SD-A
(5) SD-P
(6) SCD-A
(7) SD
(8) MOD
(9) FL
(10) SD
(11) NP+HW
(12) NP
(13) IMP-Sym
(14) IMP-Dyn
(15) IMP-GA

Figure D.8. Funnel plot of intervention vs. control for improving academic performance (at post-treatment)

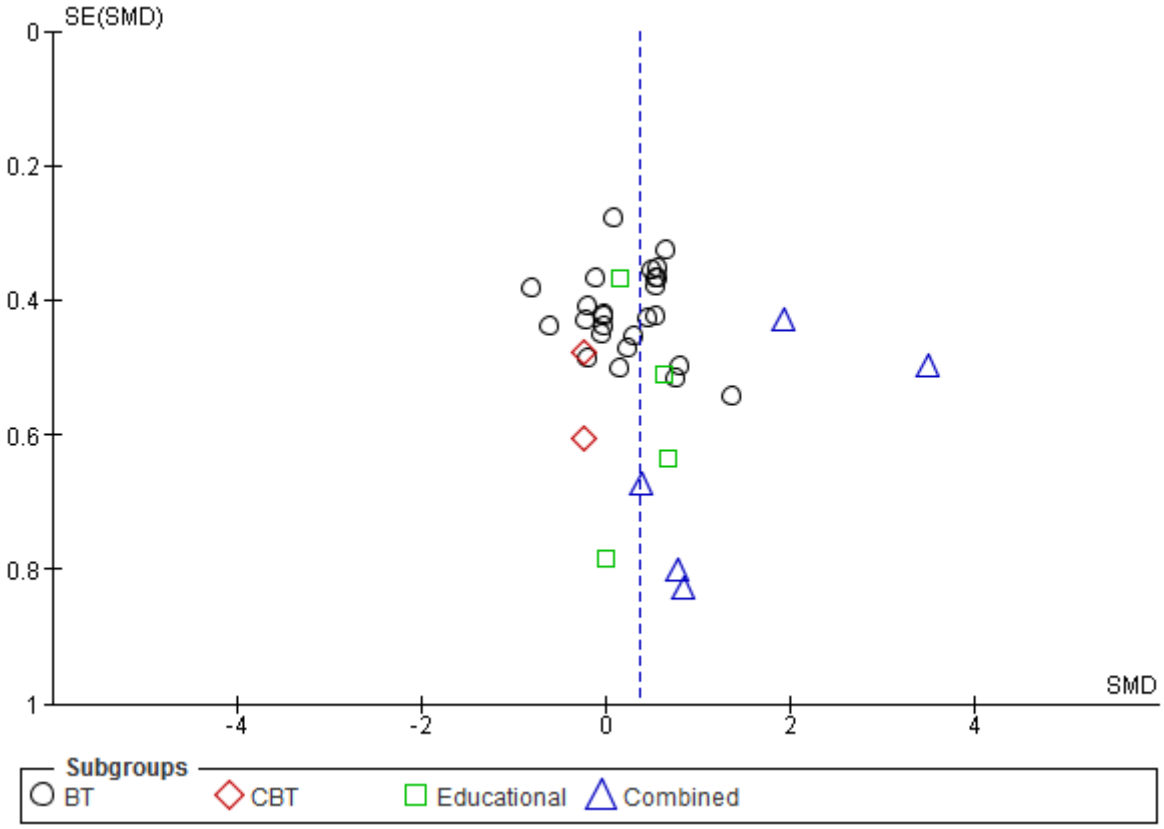
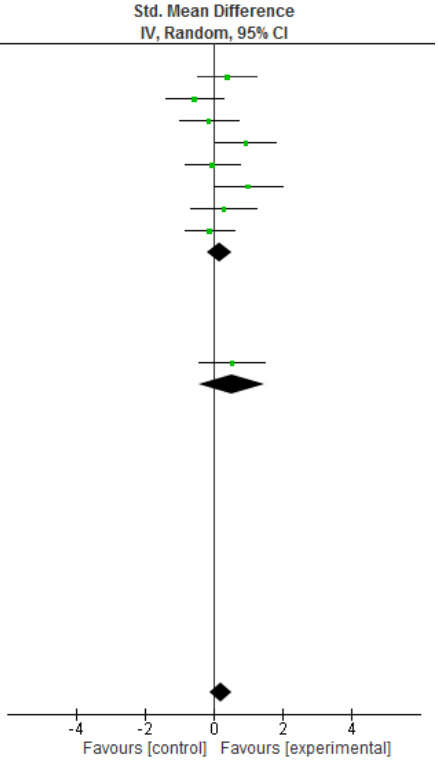


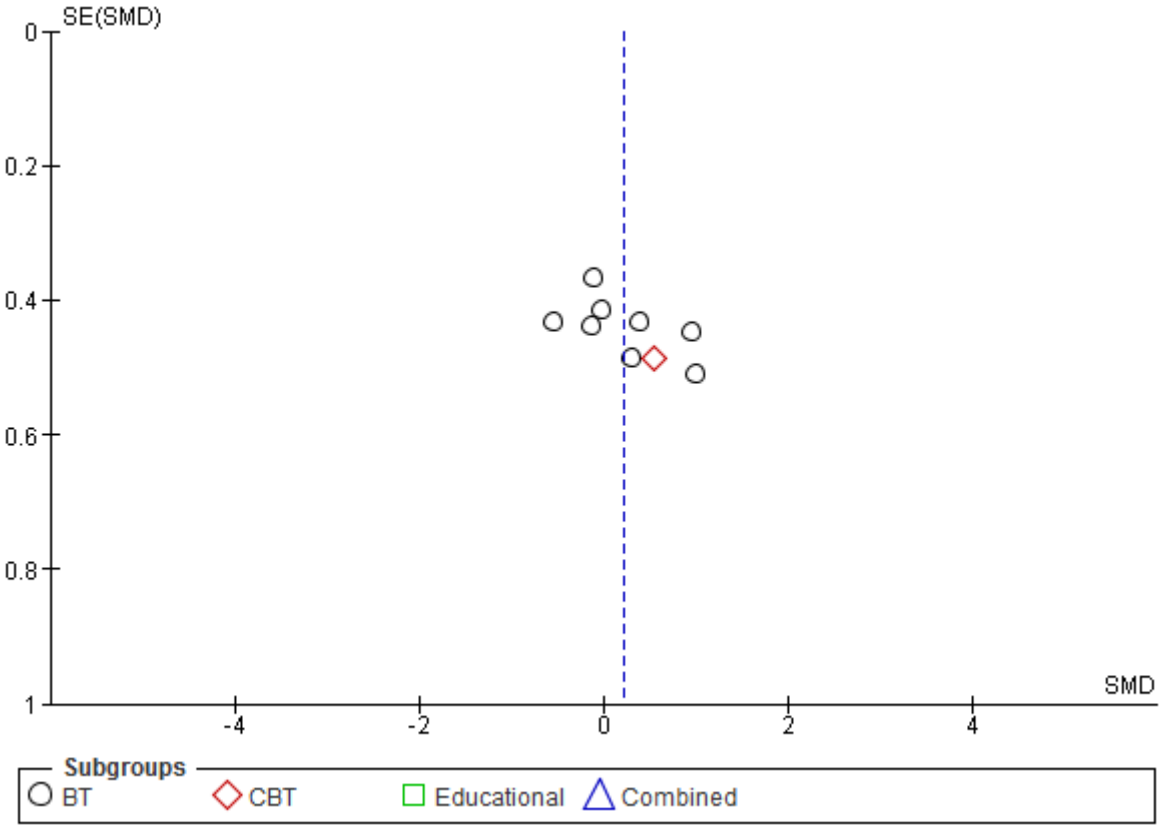
Figure C.9. Forest plot of intervention vs. control for improving academic performance (at follow-up)

Study or Subgroup	Experimental Mean	Experimental SD	Experimental Total	Control Mean	Control SD	Control Total	Weight	Std. Mean Difference IV, Random, 95% CI
5.1.1 BT								
Deffenbacher 1979a (1)	2.94	0.45	17	2.73	0.63	8	11.4%	0.40 [-0.45, 1.25]
Deffenbacher 1979a (2)	2.33	0.74	15	2.73	0.63	9	11.5%	-0.55 [-1.39, 0.29]
Denney 1977 (3)	2.53	0.94	11	2.64	0.62	10	11.3%	-0.13 [-0.99, 0.73]
Denney 1977 (4)	3.13	0.38	13	2.64	0.62	10	10.9%	0.95 [0.07, 1.83]
Denney 1977 (5)	2.63	0.8	14	2.64	0.62	10	12.2%	-0.01 [-0.82, 0.80]
Denney 1977 (6)	3.29	0.61	9	2.64	0.62	9	8.9%	1.01 [0.01, 2.00]
Holahan 1979	2.95	0.66	15	2.73	0.73	6	9.6%	0.31 [-0.64, 1.26]
Sapp 1996	60.9	7.37	15	61.6	6.12	15	14.6%	-0.10 [-0.82, 0.62]
Subtotal (95% CI)			**109**			**77**	**90.4%**	**0.20 [-0.17, 0.56]**
Heterogeneity: Tau² = 0.08; Chi² = 10.11, df = 7 (P = 0.18); I² = 31%								
Test for overall effect: Z = 1.05 (P = 0.29)								
5.1.2 CBT								
Holahan 1979	3.08	0.54	12	2.73	0.73	7	9.6%	0.54 [-0.41, 1.50]
Subtotal (95% CI)			**12**			**7**	**9.6%**	**0.54 [-0.41, 1.50]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.12 (P = 0.26)								
5.1.3 Educational								
Subtotal (95% CI)			**0**			**0**		**Not estimable**
Heterogeneity: Not applicable								
Test for overall effect: Not applicable								
5.1.4 Combined								
Subtotal (95% CI)			**0**			**0**		**Not estimable**
Heterogeneity: Not applicable								
Test for overall effect: Not applicable								
Total (95% CI)			**121**			**84**	**100.0%**	**0.23 [-0.11, 0.56]**
Heterogeneity: Tau² = 0.06; Chi² = 10.63, df = 8 (P = 0.22); I² = 25%								
Test for overall effect: Z = 1.33 (P = 0.18)								
Test for subgroup differences: Chi² = 0.45, df = 1 (P = 0.50), I² = 0%								



Footnotes
(1) SCR
(2) SD
(3) SD-P
(4) SD-A
(5) SCD-P
(6) SCD-A

Figure D.10. Funnel plot of intervention vs. control for improving academic performance (at follow-up)



Appendix E

Summary of main findings using the GRADE Working Group grades of evidence

Table E.1. Summary of main findings for reducing test anxiety at post-treatment, using GRADE Working Group grades of evidence

Outcome	Intervention	Quality assessment							Summary of findings				Comments
		No. of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other**	**No of participants**		**SMD (95% CI)**	**Quality (GRADE)**	
									Active	Control			
Reduction in test anxiety (at post-treatment) [1]	Overall	44	RCT	-	-	-	-	-	1362	847	-0.76 (-0.93, -0.58)	-	Publication bias found.
	BT	34	RCT	Serious [a]	Serious [b]	Not serious	Not serious	None	771	565	-0.83 (-1.06, -0.60)	⊕⊕○○ **Low**	
	CBT	10	RCT	Serious [a]	Not serious	Not serious	Serious [c]	None	301	156	-0.58 (-0.83, -0.33)	⊕⊕○○ **Low**	
	SST	10	RCT	Serious [a]	Not serious	Not serious	Serious [c]	None	167	93	0.02 (-0.25, 0.28)	⊕⊕○○ **Low**	
	Comb.	4	RCT	Serious [a]	Not serious	Not serious	Serious [c]	None	123	33	-1.38 (-1.96, -0.81)	⊕⊕○○ **Low**	

Table E.2. Summary of main findings for reducing test anxiety at follow-up, using GRADE Working Group grades of evidence

Outcome	Intervention	Quality assessment							Summary of findings				Comments
		No. of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other	No of participants		SMD (95% CI)	Quality (GRADE)	
									Active	Control			
Reduction in test anxiety (at follow-up) [1]	Overall	12	RCT	-	-	-	-	-	294	239	-0.87 (-1.06, -0.68)	-	
	BT	10	RCT	Serious [a]	Not serious	Not serious	Serious [c]	None	179	181	-1.15 (-1.38, -0.91)	⊕⊕○○ **Low**	Follow-up: M = 9.4 (SD = 6.4)
	CBT	3	RCT	Serious [a]	Not serious	Not serious	Serious [c]	None	107	54	-0.31 (-0.64, 0.02)	⊕⊕○○ **Low**	
	SST	1	RCT	-	-	-	-	-	8	4	-0.95 (-2.24, 0.34)	-	
	Comb.	NA	NA	-	-	-	-	-	NA	NA	NA	-	

Table E.3. Summary of main findings for reducing 'in situ' (state) test anxiety, using GRADE Working Group grades of evidence.

Outcome	Intervention	Quality assessment							Summary of findings				Comments
		No. of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other**	**No of participants**		**SMD (95% CI)**	**Quality (GRADE)**	
									Active	Control			
Reduction in in-situ (state) test anxiety [1]	Overall	3	RCT	-	-	-	-	-	128	36	-1.20 (-1.66, -0.75)	-	
	BT	3	RCT	Serious [a]	Not serious	Not serious	Serious [c]	None	72	22	-1.52 (-2.06, -0.97)	⊕○○○ **V. Low**	
	CBT	-	-	-	-	-	-	-	-	-	-	-	
	SST	2	RCT	Serious [a]	Not serious	Not serious	Serious [c]	None	29	8	-0.73 (-2.57, 1.12)	⊕○○○ **V. Low**	
	Comb.	1	RCT	-	-	-	-	-	27	6	-0.85 (-1.77, 0.08)	-	

Table E.4. Summary of main findings for improving academic performance at post-treatment, using GRADE Working Group grades of evidence

Outcome	Intervention	Quality assessment							Summary of findings				Comments
		No. of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other	No of participants		SMD (95% CI)	Quality (GRADE)	
									Active	Control			
Improvement in academic performance (post-treatment) [2]	Overall	17	RCT	-	-	-	-	-	548	363	0.37 (0.14, 0.61)	-	
	BT	14	RCT	Serious [a]	Not serious	Not serious	Serious [c]	None	379	299	0.22 (0.05, 0.40)	⊕⊕○○ **Low**	
	CBT	2	RCT	Serious [a]	Not serious	Not serious	Serious [c]	-	21	11	-0.24 (-0.98, 0.49)	⊕⊕○○ **Low**	
	SST	4	RCT	Serious [a]	Not serious	Not serious	Serious [c]	None	57	23	0.34 (-0.16, 0.84)	⊕⊕○○ **Low**	
	Comb.	2	RCT	Serious [a]	Serious [b]	Not serious	Serious [c]	None	91	30	1.58 (0.41, 2.76)	⊕○○○ **V. low**	

Table E.5. Summary of main findings for improving academic performance at follow-up, using GRADE Working Group grades of evidence

Outcome	Intervention	Quality assessment							Summary of findings				Comments
		No. of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other	No of participants		SMD (95% CI)	Quality (GRADE)	
									Active	Control			
Improvement in academic performance (follow-up) [2]	Overall	4	RCT	-	-	-	-	-	121	84	0.23 (-0.11, 0.56)	-	
	BT	4	RCT	Serious [a]	Not serious	Not serious	Serious [c]	None	109	77	0.20 (-0.17, -0.56)	⊕○○○ **V. low**	
	CBT	1	RCT	-	-	-	-	-	12	7	0.54 (-0.41, 1.50)	-	
	SST	NA	NA	-	-	-	-	-	NA	NA	NA	-	
	Comb.	NA	NA	-	-	-	-	-	NA	NA	NA	-	

Key.

GRADE Working Group grades of evidence:

High quality: Further research is unlikely to change our confidence in the estimate of effect.

Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.

Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.

Very low quality: We are very uncertain about the estimate.

Notes

1. Test anxiety was assessed using self-report measures
2. Academic performance assessed via either Grade Point Average (GPA) or examination performance

a = Risk of bias indicated (nearly 80% of items scored as unclear)

b = Substantive heterogeneity indicated ($I^2 \geq 50\%$)

c = Imprecise estimate (wide 95% confidence intervals, ≥ 0.25) or down-rated owing to lack of trials (≤ 10) and/or participations (≤ 500 in active or control)

BT = Behaviour Therapy; Comb = Combined intervention, consisting of psychotherapy (BT or CBT) and ED; CBT = Cognitive-Behavioural Therapy; ED = Educational Intervention; NA = Not Available; RCT = Randomized Controlled Trial; SA = Suspected strong association between effect size estimate and publication bias.

Appendix F

Factor loadings for EFA of MCQ-30 and IUS-12 in trait and state test anxiety datasets

Table F.1. MCQ-30 published scale structure and Geomin rotated factor loadings from EFA of trait and datasets

		EFA loadings for trait data					EFA loadings for state data				
MCQ-30 published scale structure and items		**F1**	**F2**	**F3**	**F4**	**F5**	**F1**	**F2**	**F3**	**F4**	**F5**
Subscale: Positive beliefs about worry											
MCQ-1	Worrying helps me to avoid problems in future	0.02	**0.64**	-0.03	0.12	0.03	0.14	**0.63**	-0.03	0.12	-0.05
MCQ-7	I need to worry in order to remain organized	0.16	**0.81**	0.04	-0.18	0.06	0.19	**0.77**	-0.02	0.03	-0.10
MCQ-10	Worrying helps me get things sorted out in my mind	0.01	**0.87**	-0.05	0.10	-0.11	0.00	**0.79**	0.00	0.06	0.02
MCQ-19	Worrying helps me cope	0.00	**0.83**	0.01	0.05	-0.10	-0.05	**0.77**	0.06	-0.03	0.05
MCQ-23	Worrying helps me solve problems	-0.17	**0.85**	-0.02	0.10	0.00	-0.11	**0.79**	0.04	0.00	0.11
MCQ-28	I need to worry in order to work well	-0.04	**0.86**	0.05	-0.07	0.05	0.03	**0.75**	0.04	-0.10	0.06
Subscale: Negative beliefs about worry											
MCQ-2	My worrying is dangerous for me	**0.50**	-0.11	0.00	0.05	0.30	**0.64**	-0.12	0.04	0.12	0.04
MCQ-4	I could make myself sick with worrying	**0.73**	-0.05	0.02	0.14	0.04	**0.79**	0.00	-0.10	0.10	0.00
MCQ-9	My worrying thoughts persist, no matter how I try to stop them	**0.93**	-0.02	0.02	0.00	-0.11	**0.86**	-0.01	0.11	-0.05	-0.01
MCQ-11	I cannot ignore my worrying thoughts	**0.83**	0.11	-0.08	0.00	0.02	**0.80**	0.06	0.07	-0.04	0.02
MCQ-15	My worrying could make me go mad	**0.45**	-0.05	0.15	0.12	0.38	**0.63**	-0.02	0.04	0.10	0.18
MCQ-21	When I start worrying, I cannot stop	**0.81**	0.05	-0.01	-0.15	0.17	**0.73**	0.06	0.06	-0.08	0.20

Subscale: Cognitive Confidence											
MCQ-8	I have little confidence in my memory for words and names	0.20	0.03	**0.78**	-0.03	-0.11	0.14	-0.01	**0.84**	0.02	-0.14
MCQ-14	My memory can mislead me at times	-0.05	-0.04	**0.69**	0.15	-0.02	0.11	0.03	**0.48**	0.17	0.00
MCQ-17	I have a poor memory	0.06	0.02	**0.86**	-0.01	-0.16	-0.01	-0.07	**0.91**	0.04	-0.07
MCQ-24	I have little confidence in my memory for places	0.03	-0.10	**0.71**	-0.03	0.10	-0.01	0.07	**0.75**	-0.08	0.17
MCQ-26	I do not trust my memory	-0.07	-0.05	**0.86**	0.06	0.11	-0.02	0.03	**0.90**	0.00	0.08
MCQ-29	I have little confidence in my memory for actions	-0.01	0.16	**0.76**	-0.03	0.15	0.06	0.07	**0.68**	0.06	0.07
Subscale: Need to control thoughts											
MCQ-6	If I did not control a worrying thought, and then it happened, it would be my fault	0.16	0.23	0.07	0.00	**0.44**	0.37	0.12	-0.08	0.09	0.35
MCQ-13	I should be in control of my thoughts all the time	-0.03	-0.05	-0.06	0.27	**0.62**	0.00	-0.04	-0.05	0.36	0.34
MCQ-20	Not being able to control my thoughts is a sign of weakness	0.09	0.05	-0.01	0.11	**0.68**	0.04	0.05	-0.02	0.16	**0.62**
MCQ-22	I will be punished for not controlling certain thoughts	0.01	0.21	0.05	-0.09	**0.76**	0.11	-0.01	0.12	-0.01	**0.72**
MCQ-25	It is bad to think certain thoughts	0.07	0.06	-0.05	-0.01	**0.62**	-0.02	0.02	0.09	0.13	**0.57**
MCQ-27	If I could not control my thoughts, I would not be able to function	0.09	0.12	0.11	0.06	0.21	0.08	0.14	0.00	-0.01	0.27
Subscale: Cognitive self-consciousness											
MCQ-3	I think a lot about my thoughts	**0.55**	0.05	0.00	**0.40**	-0.03	**0.42**	0.01	-0.03	**0.55**	-0.10

MCQ-5	I am aware of the way my mind works when I am working through a problem	0.09	0.05	-0.08	**0.65**	-0.20	-0.02	0.10	-0.18	**0.54**	-0.07
MCQ-12	I monitor my thoughts	-0.06	0.01	-0.07	**0.71**	0.06	0.01	0.01	-0.04	**0.62**	0.12
MCQ-16	I am constantly aware of my thinking	0.28	-0.01	0.03	**0.68**	0.11	0.22	-0.03	0.06	**0.63**	0.12
MCQ-18	I pay close attention to the way my mind works	-0.04	0.09	0.04	**0.84**	-0.04	-0.18	0.05	0.13	**0.82**	0.03
MCQ-30	I constantly examine my thoughts	0.18	0.01	0.06	**0.74**	0.10	0.09	0.00	0.10	**0.75**	0.04

Notes.

F1 = ‘Negative beliefs about worry’, F2 = ‘Positive beliefs about worry’, F3 = ‘Cognitive confidence’, F4 = ‘Cognitive self-consciousness’, F5 = ‘Need for control over thoughts’; Bold = loading ≥ 0.40; Underline = highest loading where item loads ≥ 0.40 on more than one factor.

Table F.2

Standardized factor loadings and exploratory factor analysis goodness-of-fit indices for measurement models of the Intolerance of Uncertainty Scale – 12 in trait test anxiety sample (n = 288)

	Unidimensional	Correlated factors model		Bifactor model		
Item	**model**	**Inhibitory**	**Prospective**	**General**	**Inhibitory**	**Prospective**
3. Uncertainty keeps me from living a full life.	.33	.35		.39	-.27	
6. When it is time to act, uncertainty paralyzes me.	.75	.79		.68	.49	
7. When I am uncertain I can't function very well.	.75	.80		.70	.44	
10. The smallest doubt can stop me from acting.	.73	.77		.69	.35	
12. I must get away from all uncertain situations.	.76	.81		.75	.24	
1. Unforeseen events upset me greatly.	.79		.81	.78		.19
2. It frustrates me not having all the information I need.	.78		.80	.72		.41

4. One should always look ahead so as to avoid surprises.	.62		.63	.57		.32
5. A small unforeseen event can spoil everything, even with the best of planning.	.76		.78	.78		.06
8. I always want to know what the future has in store for me.	.77		.81	.76		.27
9. I can't stand being taken by surprise.	.78		.80	.73		.33
11. I should be able to organize everything in advance.	.68		.70	.64		.29
Model fit statistics						
χ^2 *(df)*	253 (54)		172 (53)		126 (42)	
CFI	.95		.97		.98	
RMSEA [90% CI]	.11 (.10 - .13)		.09 (.07 - .10)		.08 (.07-.10)	
WRMR	.05		.04		.03	
Bifactor model ancillary statistics						
% Explained Common Variance (ECV)	-	-	-	0.82	0.08	0.10

% Omega hierarchical	-	-	-	0.87	0.11	0.11
Construct reliability (*H*)	-	-	-	0.92	0.13	0.45
Factor Determinacy (*FD*)	-	-	-	0.94	0.62	0.74

Note. CFI = Comparative Fit Index; RMSEA = Root mean square error of approximation; WRMR = Weighted Root Mean square Residual.

Table F.3

Standardized factor loadings and exploratory factor analysis goodness-of-fit indices for measurement models of the Intolerance of Uncertainty Scale – 12 in state test anxiety sample (n = 463)

	Unidimensional	Correlated factors model		Bifactor model		
Item	**model**	**Inhibitory**	**Prospective**	**General**	**Inhibitory**	**Prospective**
3. Uncertainty keeps me from living a full life.	.65	.70		.69	.20	
6. When it is time to act, uncertainty paralyzes me.	.79	.77		.74	.47	
7. When I am uncertain I can't function very well.	.80	.77		.74	.42	
10. The smallest doubt can stop me from acting.	.78	.74		.74	.19	
12. I must get away from all uncertain situations.	.67	.79		.80	-.02	
1. Unforeseen events upset me greatly.	.30		.72	.77		.18

2. It frustrates me not having all the information I need.	.72	.64	.62	.37
4. One should always look ahead so as to avoid surprises.	.70	.66	.60	.27
5. A small unforeseen event can spoil everything, even with the best of planning.	.67	.78	.81	.12
8. I always want to know what the future has in store for me.	.80	.76	.70	.12
9. I can't stand being taken by surprise.	.77	.80	.79	-.09
11. I should be able to organize everything in advance.	.73	.86	.61	-.04
Model fit statistics				
χ^2 (df)	465 (54)	193 (43)	127 (33)	
CFI	.93	.98	.99	

RMSEA [90% CI]	.13 (.12 - .14)	.09 (.08 - .10)			.07 (.06-.09)	
WRMR	.06	.04			.03	
Bifactor model ancillary statistics						
% Explained Common variance (ECV)	-	-	-	0.86	0.04	0.11
% Omega hierarchical	-	-	-	0.94	0.00	0.00
Construct reliability (*H*)	-	-	-	0.94	0.22	0.49
Factor Determinacy (*FD*)	-	-	-	0.97	0.64	0.85

Note. CFI = Comparative Fit Index; RMSEA = Root mean square error of approximation; WRMR = Weighted Root Mean square Residual.

Appendix G

Ethical approval letters from university ethics committee

Figure G.1

Ethical approval letter (ethics reference: 201605175)

Dr Peter Fisher
Psychological Sciences
University of Liverpool

School of Medicine

Faculty of Health and Life Sciences
Cedar House
Ashton Street
Liverpool
L69 3GE

T: 0151 795 4356
F: 0151 794 8763
W: www.liv.ac.uk

2nd June 2016

Dear Dr Fisher

I am pleased to say we are able to approve the request for ethical approval to undertake your research project. This is on condition that you abide by the requirements of the Human Tissue Act, and any other primary legislation that may govern the project.

Ref:	201605175
PI/Supervisor:	Peter Fisher
Title:	The relationships between metacognitive beliefs, intolerance of uncertainty, test anxiety, and examination performance (Anxiety, beliefs, and exam performance)
First Reviewer:	Jayne Garner
Second Reviewer:	David Taylor
Date of Approval:	2nd June 2016

The application was APPROVED subject to the following conditions:

Conditions

1 All serious adverse events must be reported to the Sub-Committee within 24 hours of their occurrence, via the Research Governance Officer (ethics@liv.ac.uk).

2 This approval applies for the duration of the research. If it is proposed to extend the duration of the study as specified in the application form, IPHS REC should be notified as follows. If it is proposed to make an amendment to the research, you should notify IP IPHS REC by following the Notice of Amendment procedure outlined at http://www.liv.ac.uk/researchethics/amendment%20procedure%209-08.doc

3 If the named PI / Supervisor leaves the employment of the University during the course of this approval, the approval will lapse. Therefore please contact the Institute's Research Ethics Office at iphsrec@liverpool.ac.uk in order to notify them of a change in PI/Supervisor.

Best wishes and good luck with the study.

Ann Furlong
Ann Furlong

ILT Ethics Review Group (Staff) Secretary

E: furlonga@liveprool.ac.uk
T: 0151 795 4358

Figure G.2

Ethical approval letter (ethics reference: 201602153)

UNIVERSITY OF LIVERPOOL

School of Medicine

Faculty of Health and Life Sciences
Cedar House
Ashton Street
Liverpool
L69 3GE

T: 0151 795 4356
F: 0151 794 8763
W: www.liv.ac.uk

Dr Peter Fisher
Psychological Sciences
Waterhouse
Liverpool,
L69 3BX

16th February 2016

Dear Peter

I am pleased to inform you that the ILT Ethics Review Group (staff) has approved your application for ethical approval. Details of the approval can be found below:

Ref:	201602153
PI/Supervisor:	Peter Fisher
Title:	Do metacognitive beliefs or intolerance of uncertainty beliefs predict test anxiety in undergraduate medical students? (Beliefs about exam anxiety)
First Reviewer:	John Smith
Second Reviewer:	Ben Shaw
Date of Approval:	16th February 2016

The application was APPROVED subject to the following conditions:

Conditions

1 All serious adverse events must be reported to the Sub-Committee within 24 hours of their occurrence, via the Research Governance Officer (ethics@liv.ac.uk).

2 This approval applies for the duration of the research. If it is proposed to extend the duration of the study as specified in the application form, IPHS REC should be notified as follows. If it is proposed to make an amendment to the research, you should notify IP IPHS REC by following the Notice of Amendment procedure outlined at http://www.liv.ac.uk/researchethics/amendment%20procedure%209-08.doc

3 If the named PI / Supervisor leaves the employment of the University during the course of this approval, the approval will lapse. Therefore please contact the Institute's Research Ethics Office at iphsrec@liverpool.ac.uk in order to notify them of a change in PI/Supervisor.

Best wishes and good luck with the study.

Ann Furlong

Ann Furlong
ILT Ethics Review Group (Staff) Secretary

E: furlonga@liveprool.ac.uk
T: 0151 795 4358