Emotional Determinants of Test Anxiety and Academic Performance

A thesis presented in partial fulfillment of the requirements for the degree of

Doctor of Clinical Psychology

At Massey University

Palmerston North, New Zealand

Edwin Chun-Hong CHIN

2014

~ To our final promise ~

~ Will miss you always ~

- 匡 -

Abstract

The effect of test anxiety on academic performance has been studied extensively throughout the past few decades. Recent developments in test anxiety research have largely been based within the cognitive psychology framework, where different components of *working memory* were identified to mediate the relationship between test anxiety and test performance. Similarly, the field of educational psychology has expanded this area of research to identify the different pathways in which emotional states can serve both activating and deactivating roles towards learning and achievement. From the clinical psychology perspective, the connection between emotional experiences and thought processes is an integral part of assessment which then informs ways of intervention. However, there is limited research that explicitly examines the relationship between general emotional distress and more specific forms of test-related distress, such as the cognitive and physiological components of test anxiety.

The Tripartite Model of Emotions (TME) is used to explore the connection between general emotional distress and test anxiety. The model proposed that the experiences of depression and anxiety are predisposed by a combination of three high-order dimensions of emotional distress: positive affect (or lack thereof), negative affect, and physiological hyperarousal. While researchers have identified these tripartite factors to be significant predictors of various health and performance outcomes, the degree to which the tripartite model may account for the experience of test anxiety, as well as the level of academic performance, remains unclear.

In the present study, 642 secondary school students (aged 16-19) completed a questionnaire comprised of measures of test anxiety, depression, anxiety, and the tripartite dimensions. This enabled a cross-sectional investigation into the validity of the tripartite model of emotions, as well as how test anxiety may be predicted by the higher-order factors of emotional distress. The grade point averages of a sub-sample of 188 students were gathered, which enabled a prospective investigation

into how these emotional variables influenced academic performance. Structural equation modeling was employed to simultaneously test the relationships among the aforementioned variables, and to identify an explanatory model for academic performance.

There was support for the tripartite factors' hypothesized influence on depression and test anxiety. Specifically, low levels of positive affect (PA) and high levels of negative affect (NA) influenced the experience of depressive symptoms, while high levels of negative affect and physiological hyperarousal (PH) influenced test anxiety symptoms. Negative affect was not revealed to have a direct influence on test performance. Rather, its influence may be mediated by more specific factors, including the cognitive and affective features of test anxiety. In the presence of test-related worries, negative affect may indirectly impair test performance. However, in the absence of such worries, there is potential for negative affect or the sense of emotional apprehension to facilitate better performance.

Acknowledgements

The completion of this thesis would not have been possible without the guidance and support of many people. I would first like to acknowledge my supervisors, Mei-Wah Williams, Joanne Taylor, and Shane Harvey. Mei, it was an absolute pleasure and privilege to have been your student. Despite being a long-distance supervisor you were never hard to reach, and your sound guidance and advice was very much appreciated. Jo, you have my greatest respect and admiration for your knowledge, efficiency, attention to detail, and above all, your dedication towards helping your students. Your contribution both as an academic supervisor and clinical mentor has been invaluable. Shane, you have been a source of stimulating conversations and intellectual inspiration from the beginning, and I cherish all that I had learnt from you throughout my postgraduate studentship.

Thanks are due to the staff at Palmerston North Girls High School, namely the student counsellor, Brenda Pomana-Whale; senior dean and teacher, Ginnie Harcombe; and principal, Melba Scott, who provided consultation and insight during the initial development of this research. My appreciation extends to all the principals of participating schools for offering their support for this project. Of course, I would like to thank all the participants for making this project possible.

I would like to thank Prof. Bruce Chorpita, Prof. Peter Muris, Prof. Jeff Laurent, and Dr. Dejan Stevanovic, for sharing their knowledge, psychometric resources, and insightful suggestions in support of this research. Thank you to my examiners, Prof. John Spicer, Prof. Gavin Brown, and Prof. Jerrell Cassady, for their critical review of this thesis, and for their guru-class advice to help this project reach its full potential.

I would also like to take this opportunity to thank those who played a role in my clinical training, which made up a substantial and valuable part of this doctoral qualification. My sincerest gratitude is reserved for my clinical supervisor Erin Mooney for her mentoring and support through my internship. As per the consensus amongst my fellow interns, she is the epitome of awesomeness!! Thanks are due to my office buddy Anne Harvey for her immense tolerance of my

questions and for providing a listening ear whenever I felt like I was hitting walls. Thanks are also due to the clinical lead, Robyn Girling-Butcher, for overseeing my internship and for her active involvement in the DClinPsych programme. I would like to thank Cheryl Woolley and Janet Leathem for their invaluable advice as mock-examiners, and for their kind words of reassurance. My gratitude extends to Dirk Badenhorst, Przemek Dawidowski, and Rody Withers, for supervising my community placement practicums, as well as to the clinicians at the Massey Psychology Clinic, who in various ways have all helped to broaden my clinical knowledge. Last but not least, I would like to express a special thanks to Senior Clinican Jan Dickson for her wisdom and tireless efforts in supporting me and my fellow interns.

Special thanks are due to my intern-buddy Jodi Field– my internship would not have been the same without having her across the corridor, always ready to offer words of reassurance during one of our many sanity breaks. I would like to thank my other fellow interns, Kara Duxfield, Raewyn Barry, and Evelyn Aranas, for their support and encouragement. My gratitude extends to Angela Callear, Elizabeth Yan, Laura Howard, Tomoko Yamaguchi, Rifshana Fathimath, Andreas Marwick, and other pre- and post-interns for their friendship.

My friends outside of the clinical programme provided a much-appreciated sanctuary where conversations did not involve psychology, research, or the various stages of pregnancy and childbirth! Personal thanks are due to Sandy, Kevin, Grace, Tara, Karen, William, Jenny, Stan, Gloria, and Victor, for their companionship and for simply being who they are.

Thanks are due to my immediate and extended family, most of whom had been wondering for the past five years why I still didn't have a job. Thank you to my dear grandma, from whom I receive an ever-present love that transcends the need for words. Thank you also to my aunty Connie for playing the aunty role so perfectly by saying exactly what I need to hear. Finally, I sincerely thank my mother and sister for their immense patience, unwavering support, and grounded guidance.

Table of Contents

List of Tables

List of Figures

List of Acronyms

AFARS	-	Affect and Arousal Scale
AIC	-	Akaike Information Criteria
APA	-	American Psychiatric Association
CDI	-	Children's Depression Inventory
CFI	-	Comparative Fit Index
CR	-	Critical Ration
df	-	Degrees of freedom
ECVI	-	Expected Cross-Validation Index
EFC	-	Emotion-focused coping
GAD	-	Generalised Anxiety Disorder
GPA	-	Grade point average
n.d.	-	No date
(N)MAR	-	(Not) missing at random
NA	-	Negative affect
NCEA	-	National Certification of Educational Achievement
OCD	-	Obsessive Compulsive Disorder
PA	-	Positive affect
PET	-	Processing efficiency theory
PFC	-	Problem-focused coping
PGFI	-	Parsimony Goodness of Fit Index
PH	-	Physiological hyperarousal
PH-PANAS-C	-	Positive and Negative Affect Schedule for Children & Physiological Hyperarousal Scale for Children
PNFI	-	Parsimony Normed Fit Index
RADS-II	-	Reynolds Adolescent Depression Scale – Second edition
RCMAS	-	Revised Children's Manifest Anxiety Scale
RMSEA	-	Root Mean Square of Approximation
RTT	-	Reactions to Tests
SCARED-R	-	Screen of Children's Anxiety Related Emotional Disorders – Revised
SD	-	Standard deviation
SEM	-	Structural equation modeling
STAI-C	-	State Trait Anxiety Inventory for Children
TAI	-	Test Anxiety Inventory
TLI	-	Tucker-Lewis Index
TME	-	Tripartite Model of Emotions

Introduction

The present research is an integration of two major areas of investigation relevant to teenagers in their senior years of secondary school education (ages 16 – 19). Firstly, academic qualifications have become the fundamental requirement for entry into the competitive job market, and students across the world are faced with increasing pressures to meet academic expectations. In some societies, the drive for high school students to maintain a competitive edge in the academic arena has facilitated the proliferation of *cram schools* and the supplementary tuition market (Bray & Kwok, 2003; Kwok, 2004; Nath, 2008). In a survey of 423 Year 12 (age 16-18) students in Australia, participants identified *studying for exams* and *examination results* as being among the most concerning stressors that they experience (Kouzma & Kennedy, 2004). While students may react to tests and exams with different emotions, a particularly prominent emotion elicited by such pressures is anxiety.

Test Anxiety, or alternatively *examination stress* or *evaluation anxiety*, has been a popular research topic over the past few decades, resulting in many different opinions regarding the definition of this construct. In the strictest sense, *anxiety* is a term used to capture people's reaction to the anticipation of future danger, as opposed to *fear* which is the response towards an existing, present or potential threat (Perkins, Kemp, & Corr, 2007; Sadock & Sadock, 2007). However, this distinction appears to be used rather interchangeably within test anxiety research, as test anxiety may be experienced before, as well as during, a testing situation (Cassady, 2004; Hunsley, 1985; Zeidner, 1998). For example, before an exam, test anxious students may negatively evaluate their own ability to tackle the exam, and perceive a potential harm to their social and academic standing as consequences of failure (Cassady, 2004; Deffenbacher, 1980; Lay, Edwards, Parker, & Endler, 1989). As students carry out the exam, the anxiety may take the form of preoccupation over a perceived lack of preparation, and attentional bias towards negative stimuli such as difficult or unexpected questions (Sarason, 1984, 1986; Vasey, El-Hag, & Daleiden, 1996). Cassady (2004) argued that the impact of test anxiety persists

even after the completion of the exam, as it influences the way in which students reflect on their performance and attribute their successes and failures (Forsyth, Story, Kelley, & McMillan, 2009).

Different conditions of academic testing have also demonstrated different levels of bias towards test anxiety. For example, dynamic assessments involving seatwork (e.g., essay writing, where teachers may provide assistance during testing) are associated with lower levels of test anxiety than the conventional end-of-year examinations (Meijer, 2001; Pintrich & de Groot, 1990). This has particular implications for the New Zealand educational context, where the conventional examination system has been replaced by the current NCEA system which resembles the dynamic assessment approach.

The experience of test anxiety has been demonstrated over the past few decades to be largely counter-productive to achievement (Cassady, 2004; Deffenbacher, 1980; Hembree, 1988; Musch & Bröder, 1999). However, despite this long-standing line of research, theoretical explanations of how test anxiety affects test performance have been under ongoing development and review. Hence, this study aims to examine the level of test anxiety that New Zealand high school students experience, and how it may affect their academic performance.

The second area of investigation surrounds how broader emotional distress may contribute to the relationship between test anxiety and academic functioning. This follows the plethora of research into a growing list of identified predictors of academic performance, including self-regulation and motivational factors (Elliot, McGregor, & Gable, 1999; Pintrich & de Groot, 1990), cognitive appraisals of control (Pekrun, 2006), self-efficacy (Chemers, Hu, & Garcia, 2001; Thelwell, Lane, & Weston, 2007), and working memory and learning processes (Curcio, Ferrara, & de Gennaro, 2006; Derakshan & Eysenck, 2009). As parents, teachers, or the students themselves, may not readily recognise these underlying cognitive mechanisms, there is impetus to establish a parsimonious model for predicting test anxiety and academic performance, using variables that are conceptually basic enough to generate practical implications. Of greater relevance to the field of clinical

psychology, classes of emotional disturbance such as anxiety and mood disorders can often encapsulate the aforementioned variables that affect learning and academic achievement. Indeed, anxiety and depression have been linked to reduced academic performance across the high school and university age groups (Fröjd et al., 2008; Hysenbegasi, Hass, & Rowland, 2005; Owens, Stevenson, Hadwin, & Norgate, 2012). This highlights the need to identify and clarify how general emotional distress (such as the tendency to feel upset) can predict more specific symptoms of depression, anxiety, and test anxiety, as well as these factors may be expected to influence educational outcomes.

The present study begins by reviewing the theoretical foundations of test anxiety, and examining how the cognitive and emotional components of test anxiety are expected to influenced test performance. The chapters then explore the apparent complexity surrounding this relationship in the context of the numerous mediating mechanisms that are identified through with fields of cognitive and educational psychology research. Following this, the study presents the rationale for examining the precursory role of a broader and more parsimonious construct of 'emotional distress'. Hence, the core research question is formulated: how do factors of general emotional distress impact on the experience of test anxiety and subsequent test performance?

In examining this question, there was a need to account for the potentially confounding effects of variables that are closely related to test anxiety. In particular, the role of depression and anxiety were deemed important because they are conceptually common with test anxiety, are the most prevalent types of emotional disturbances experienced by the teenagers, and have been shown to influence academic performance. Hence, these factors were also included in the present study. The Tripartite Model of Emotion (Clark & Watson, 1991) provided the ideal theoretical framework for examining the research question while controlling for the effects of depression and anxiety.

Chapter One Test Anxiety

1.1 Theoretical Foundations of Test Anxiety

The concept of anxiety and neuroticism attracted much research interest during the rise and expansion of experimental psychology research in the early 20th century. During that time, researchers sought to examine the physiological reactions associated with anxiety by inducing the emotional state in laboratory settings, such as introducing electric shocks or suddenly tilting participants backwards on their chair (Blatz, 1925; Landis, 1924). Acknowledging the artificiality of such research, C. H. Brown (1938) sought to study the concept of anxiety through more naturalistic means by focusing on the effect of academic examinations – an already well-identified trigger of distress. C. H. Brown hence examined emotional reactions toward tests from an anxiety framework by asking participants to rate how much they experienced particular symptoms before taking a test. This study provided the first investigation into the physiological (e.g., dry throat, palpitations) as well as subjective psychological feelings (e.g., irritable, worry) associated with test-taking, and initiated a proliferation of research into how psychological factors may influence test performance.

A particularly prominent theoretical framework used to study test anxiety was the stimulus-response (S-R) theory. This theory posits that the experience of anxious states elicits a variety of physiological and cognitive responses, which may be competing (task-irrelevant) or facilitating (task-relevant) in nature (Nicholson, 1958). Task-irrelevant responses such as feelings of inadequacy, helplessness, heightened somatic reactions, and anticipations of punishment interfere with task completion and reduce task performance (Mandler & Sarason, 1952). Grounded within the classical conditioning framework, S-R theory posits that the elicited response would depend on the aversive nature of the stimulus, such as when a highly difficult test elicits greater levels of anxiety compared to an easy test (Nicholson, 1958).

Another prominent theory that was influential to the study of test anxiety was drive theory, which focuses on the role and strength of internal drive states. This theory postulates that people's behavioural responses are guided by the need to satisfy intrinsic needs, such as hunger, thirst, or the avoidance of pain (Wolpe, 1950). When applied to test anxiety research, it posits that the testing situation evokes the drive to satisfy the requirements of the task, and elicit task-oriented behaviours that serve to help complete the task and reduce the drive (Mandler & Sarason, 1952). Similar to the S-R theory, drive theory acknowledges the role of task-irrelevant responses in impairing task performance, but views such responses as being part of an innate and readily evoked repertoire of responses (Nicholson, 1958). Mandler and Sarason's (1952) study was the first to incorporate S-R and drive theories to examine the relationship between test anxiety and test performance. Participants were asked to rate various subjective experiences before and during a testing condition, including feelings of uneasiness, accelerated heartbeat, perspiration, emotional interference, worry, as well as their attitudes (e.g., likes and dislikes) towards tests. This study also marked the development of the first self-report measurement tool for test anxiety. Since then, there was progression in the study of test anxiety throughout the subsequent six decades, particularly in the development of its theoretical conceptualisation.

1.2 Contemporary Models and Measures of Test Anxiety

An influential theoretical shift was the expansion of Mandler and Sarason's (1952) singular unidimensional view of test anxiety into multi-dimensional models, which was also closely accompanied by the development of new measures of test anxiety. Factor analytic studies carried out with these measures in turn perpetuated theoretical development in this area by providing insight into better defining the components of test anxiety. Mandler and Sarason (1952) initially viewed test anxiety as a singular unidimensional construct that was activated when confronted by a challenging testing situation, and their Test Anxiety Questionnaire reflected this unidimensional conceptualisation (Sarason & Mandler, 1952). However, the authors acknowledged that test anxiety consisted of a combination of physiological,

cognitive, and motivational components, and researchers soon began to expand on the concept and develop multi-dimensional models of test anxiety. A particularly influential model was the two-factor worry-emotionality bidimensional model (Liebert & Morris, 1967). Briefly, *emotionality* represents the awareness of the physiological manifestations of anxiety, including increased heart rate and palpitations, stomach aches and nausea, dizziness, and feelings of panic (Deffenbacher, 1980; Hembree, 1988). *Worry,* on the other hand, is the cognitive expression of concern about an evaluative situation, and may include worries about potential failure and its consequences, feeling unprepared for the test, and having self-oriented doubts over one's own capabilities (Deffenbacher, 1980; Sarason, 1984; Spielberger & Vagg, 1995; Wine, 1971). Based on this model, Liebert and Morris (1967) established the Worry-Emotionality Questionnaire (WEQ) to capture these components of test anxiety.

The differentiation of test anxiety from other forms of anxiety disorders may be difficult and at times arbitrary, and it soon became apparent that there is considerable convergence between *test anxiety* and the broader construct of *anxiety*. For example, an individual's exclusive and consistent anxiety towards evaluative situations may be due to a fear of embarrassment, which is subsumed under the symptoms of social phobia according to the DSM-IV-TR (American Psychiatric Association, 2000). Some may experience test anxiety in conjunction with either another anxiety disorder, such as generalised anxiety disorder, or as a product of a trait anxiety predisposition. For example, King, Mietz, Tinney, and Ollendick (1995) identified a significant difference in self-reported general and trait anxiety between adolescents with low and high levels of test anxiety. Similarly, Beidel, Turner, and Trager (1994) identified high comorbidity between test and trait anxiety for a sample of elementary students.

Drawing on the widely accepted distinction between state and trait anxiety, Spielberger (1972) highlighted the difference between state and trait *test* anxiety. Specifically, state test anxiety represents the experience of the aforementioned worry and emotionality symptoms elicited during the testing situation. Trait test anxiety, on the other hand, represents the dispositional tendency for the individual

to experience apprehension specifically towards evaluative situations, and to hold expectations of negative outcomes (Elliot et al., 1999). The conceptualisation of test anxiety as a situation-specific anxiety trait formed the theoretical basis of the Test Anxiety Inventory (TAI; Spielberger, Gonzales, Taylor, Algaze, and Anton, 1978), which became widely used and subsequently provided support for the applicability of the worry-emotionality model across age, gender, and cultural groups (Chapell et al., 2005; Rocklin & Ren-Min, 1989; Sharma & Sud, 1990).

While there was support for the TAI's construct validity and factor structure, it was commonly observed that the worry and emotionality dimensions, as subscales and as latent variables, were highly correlated (Benson & El-Zahhar, 1994; Hodapp & Benson, 1997), suggesting a high degree of overlap between how worry and emotionality were defined, hence blurring the distinction between the two dimensions. In response to the critique that existing two-dimensional models lack the sensitivity for capturing the distinct facets of test anxiety, Sarason (1984) later separated the construct of emotionality into *tension* and *bodily reactions.* The former captures the general feeling of nervousness such as *feeling panicky, freezing up, or feeling jittery* during tests. The latter captures more specific bodily sensations including *having headaches, dry mouth, heart beating faster, and stomach upsets.* As for the cognitive components, he separated the construct of worry into *worry* and *task-irrelevant thinking.* The former covers largely the same types of worries that are captured by previous models, involving high threat appraisal of the test situation and interpersonal comparisons. The latter, on the other hand, is not related to achievement, but simply covers the manifestations of being distracted or unable to focus, such as *thinking about recent past events, thinking about being somewhere else, having fantasies,* and *daydreaming* during tests. The resulting four-factor model is represented in the Reaction to Tests (RTT; Sarason, 1984) questionnaire, which contains four subscales to measure each corresponding factor.

Subsequent developments of test anxiety measures have been characterised by further refinement of existing measures, which were predominantly based on the worry-emotionality dimensions. For example, the Revised Test Anxiety Scale (RTA;

Benson & El-Zahhar, 1994) was developed by reducing redundant items from the RTT and combining those with several items from the TAI. Similarly, the Cognitive Test Anxiety Scale, which focuses only on assessing the cognitive components of test anxiety, was validated through comparison with existing measures such as the RTT, TAI, and RTA (Cassady & Johnson, 2002).

While the theoretical development of the test anxiety construct was heavily restricted within the worry and emotionality dimensions, development of psychometric measures of test anxiety has extended beyond the cognitive and physiological domains. For example, a Revised Worry Emotionality Scale (Meijer, 2001) incorporated items of self-confidence in addition to the existing cognitive and physiological components. Others urged the need to consider interpersonal concerns such as the appraisal of social derogation, which was explicitly examined in the FRIEDBEN Test Anxiety Scale (FTA; Friedman & Bendas-Jacob, 1997). The recently developed Test Anxiety Inventory for Adolescents and Children (TAIAC; Lowe et al., 2008) incorporates the role of facilitative anxiety as one of its factors, providing a contemporary measure of this relatively neglected topic in test anxiety research.

1.3 Interventions for Test Anxiety

An important implication of developing theoretical models of test anxiety is to improve our understanding into the mechanisms through which psychological interventions may elicit positive change. Two meta-analyses summarise the literature into the efficacy of test anxiety interventions within the past three decades (Ergene, 2003; von der Embse, Barterian, & Segool, 2013). Ergene (2003) identified an overall medium to large effect size in reductions of test anxiety symptoms, with the most promising intervention being the combination of skill-focussed training with behavioural and cognitive techniques. Systematic desensitisation was identified as the behavioural technique with the highest effect size, while cognitive structuring was the superior cognitive technique (Ergene, 2003). These findings were largely replicated by von der Embse et al. (2013) who reviewed more recent research conducted in the past decade. The authors pointed

out a relatively new area of research into the use of biofeedback in helping students monitor their physiological responses, with promising results regarding the reduction of somatic anxiety symptoms (Bradley et al., 2010; von der Embse et al., 2013). It is apparent that interventions for test anxiety typically belong within the cognitive-behavioural framework, and thus largely demonstrate the cognitive and behavioural mechanisms of change. In comparison, there is a scarcity of research into the mechanisms of affective states or emotional distress in influencing the experience of test anxiety.

The concept of emotions as the target of intervention for test anxiety has been studied within the literature on coping. For example, Struthers, Perry, and Menec (2000) compared how emotion-focussed coping (EFC) and problem-focussed coping (PFC) mediated the relationship between exam stress and academic functioning. The results suggested that promoting PFC (e.g., 'I make a plan of action', 'I buy a study guide', or 'I do what has to be done one step at a time') would serve to reduce the impact of exam stress on academic performance. However, a characteristic of the literature on coping is that EFC styles are partly conceptualised from a maladaptive standpoint, with strategies such as minimalising the problem, self-blame, avoidance, and wishful-thinking (e.g., Carver & Connor-Smith, 2010; Piko, 2001). Indeed, Struther et al.'s measure of EFC included both adaptive factors of general emotional support ('I discuss my feelings with someone') as well as maladaptive factors of denial ('I refuse to believe that it happened') and academic disengagement ('I skip class', 'I give up trying to reach my goals'). Hence it is not surprising that strategies surrounding emotion regulation were not found to be conducive to better outcomes.

Within clinical psychology practice, a person's mood or recent emotional experiences are often the first line of investigation that is carried out, and the person's ability to adapt to negative emotions is a key therapeutic indicator. Thus, it is useful to further understand how such emotional experiences are expected to predict various outcomes in order to establish an informed prognosis. As such, clarification regarding the precursory role of general emotional distress on the experience of test anxiety is warranted. Based on the identified deleterious effects

of test anxiety on test performance, such an investigation will also help identify how general emotional distress may influence academic functioning.

1.4 Test Anxiety and Academic Performance

Researchers have long acknowledged that anxiety may play a positive role in improving task performance, such as in enhancing executive functioning, initiating goal-oriented behaviours, or triggering internal drives to achieve (Alpert & Haber, 1960; Elliot et al., 1999; Mandler & Sarason, 1952). A threshold effect of test anxiety with performance has been replicated in research across student age groups, suggesting that an optimally mild level of anxiety can promote good outcomes (Anderson & Sauser, 1995; Fletcher, Lovatt, & Baldry, 1997; Gregor, 2005; Owens et al., 2008). However, the facilitative role of test anxiety has received relatively little attention in research over the past few decades (Anderson & Sauser, 1995; Martin & Marsh, 2003), as most of the theoretical development in this area has focused on the detrimental effects of anxiety and how it impairs performance (e.g., Cassady & Johnson, 2002; Tobias, 1985; Sarason, 1984).

There is a consistently replicated finding that a higher level of test anxiety is related to poorer performance on academic examinations (e.g., Cassady, 2004; Gregor, 2005; Hunsley, 1985). While this relationship is often identified across different modes of research, the strength of the effect appears to be small. For example, a comprehensive meta-analysis of 562 studies identified that test anxiety has a weak negative correlation with various indicators of academic performance (Hembree, 1988). A review of the literature which used measures of both the worry and emotionality components of test anxiety (e.g., the TAI) reveal that it accounts for about 5% of the variance in exam performance, and that the effect size is generally small (Chapell et al., 2005; Hembree, 1988; Musch & Bröder, 1999). Further, there are apparent gender differences in both the level of anxiety experienced and its magnitude of influence on test performance.

It is often found that females experience higher levels of test anxiety than males – a difference that is pervasive across early childhood education to university (Cassady & Johnson, 2002; Chapell et al., 2005; Devine, Fawcett, Szűcs, & Dowker, 2012;

Hembree, 1988; Sharma & Sud, 1990). However, there is inconsistency in the literature on gender differences regarding the magnitude of the relationship between test anxiety and test performance. For example, Chapell et al. (2005) found statistically significant differences in grade point average between high-anxious and low-anxious female college students but not males. Conversely, Miller and Bichsel (2004) identified a significant relationship between anxiety and maths performance for male college students but not females. These inconsistencies highlight the complexities in the potential underlying mediating factors that influence the connection between emotional states and achievement. For example, Pekrun's (2006) Control-Value theory highlighted how the cognitive appraisal of a testing situation, such as whether it is threatening or whether the individual can exert control over the outcome, determines whether the resulting emotional experience is activating or deactivating, serving to promote or reduce performance, respectively. This is consistent with Bandalos, Yates, and Thorndike-Christ's (1995) finding that perceived self-efficacy and patterns of success / failure attributions explained the gender differences in levels of test anxiety. Pekrun, Frenzel, Goetz, and Perry (2007) also asserted that while the structural relationship between cognitive appraisals and achievement emotions are universal across gender, there are gender differences in the actual type of appraisal and the resulting emotional experiences. For example, females may have higher levels of perceived threat towards testing situations than males (Cassady & Johnson, 2002).

It is difficult to then make inferences regarding how gender differences in appraisal processes influence the relationship between emotions and performance. On the one hand, negative appraisals can impair performance through the resulting deactivating emotions such as hopelessness and boredom, as well as through the development of performance-avoidance goals (Daniels et al., 2009; Pekrun et al., 2007). Hence, if it is argued that negative appraisals are more prevalent for females, then it stands to reason that there is a pertinent connection between appraisal processes and test performance for females. On the other hand, the argument that negative appraisals are the product of gender-based tendencies also raises the argument that appraisals are independent from performance. For

example, Ma and Xu (2004) proposed that there are differences in the way males and females reflect on past test performance, insofar as males may develop anxiety following poor performance, whereas females' anxiety are more intrinsic, stable, and independent from past performance. From this perspective, the connection between appraisal processes and test performance is less substantial than for males. Given such complexities, there is a need for research to account for the potential gender effects within this topic of research.

A finding that is more consistent within test anxiety research is that only the cognitive or worry aspect of anxiety is significantly associated with lower academic performance (Deffenbacher, 1980; Hembree, 1988). Studies have generally found that emotionality has no significant relationship with test performance (Chapell et al., 2005; Deffenbacher, 1980; Hembree, 1988), or that it is associated with impaired performance only when there are also high levels of worry (Morris, Davis, & Hutchings., 1981). These findings have largely been interpreted from two main theoretical standpoints: interference and deficit models.

The finding that only the worry component has a deleterious effect on performance is consistent with the *interference* view on the relationship between test anxiety and test performance. Interference models of test anxiety focus on the intrusion of thoughts during a test situation, placing greater demands on the individual's attention capacity, and thus reducing the cognitive resources available to complete the task (Wine, 1980). Similar arguments were adopted in Sarason's (1984) model of *cognitive interference,* which focuses on the intrusive internal dialogue that people experience during tests, such as being preoccupied with doubts over one's own capabilities. In addition to self-defeating thoughts, cognitive interference may also manifest as 'freezing up', daydreaming, or otherwise engaging in task-irrelevant thinking such as thinking about current events (Musch & Bröder, 1999; Sarason, 1984). Hence, the interference framework provides an account of why the cognitive *worry* component of test anxiety exerts a much greater effect on academic performance than that of emotionality.

Interference models carry the implicit assumption that worry mediates the impairing influence of anxiety on task performance, regardless of pre-existing factors such as level of ability, knowledge, or study skills. Indeed, students across ability levels may experience cognitive interference and have problems retrieving previously learned information only during an examination (Naveh-Benjamin, McKeachie, & Lin, 1987; Zohar, 1998). However, the predicted group differences in cognitive task performance between high-anxious and low-anxious groups have not always been observed (e.g., Blankstein, Flett, Boase, & Toner, 1990; Blankstein, Toner, & Flett, 1989). Further, the reduction of worry alone is not reliably linked to improved performance (Allen, Elias, & Zlotlow, 1980; Gregor, 2005), with other studies identifying that worry accounts for only approximately 5% of the variance in task performance (Cassady & Johnson, 2002; Chapell et al., 2005; Hembree, 1988; Musch & Bröder, 1999).

Deficit models consider factors beyond the effects of cognitive interference that occurs during the test-taking situation, and focus instead on the processes involved during learning and test preparation. The central premise of deficit models is that individuals become test anxious due to known deficiencies in study skills and test-taking abilities (Naveh-Benjamin, McKeachie, Lin, & Holinger, 1981; Tobias, 1985). This follows the observation that students with high test anxiety tend to have less effective study skills, particularly around the processing and retrieval of learning materials (Kirkland & Hollandsworth, 1979, 1980; Kleijn, van der Ploeg, & Topman, 1994).

As such, deficit model researchers have examined the predictors of performance through the cognitive psychology framework (Aronen, Vuontela, Steenari, Salmi, & Carlson, 2005; Derakshan & Eysenck, 2009; Eysenck & Calvo, 1992). In particular, cognitive theories surrounding working memory (Baddeley, 1996, 2001) have been adopted to examine the applicability of deficit models, such as information processing theory (Humphreys & Revelle, 1984), processing efficiency theory (PET; Eysenck & Calvo, 1992), and most recently, attentional control theory (Eysenck, Derakshan, Santos, & Calvo, 2007). The progression in development of these theories is characterised by the increasingly specific identification of mechanisms of

working memory that impair different types of cognitive demands (Derakshan & Eysenck, 2009). For example, Owens et al.'s (2008) examination of PET identified how *verbal* working memory mediated the relationship between trait anxiety and academic performance while *spatial* working memory did not. Based on attentional control theory, Derakshan and Eysenck (2009) distinguished between the flexibility and inhibitory roles of executive functioning to explain how the influence of anxiety may differ across different cognitive demands. Further studies have adapted such explanations and have identified how specific components of executive functioning may have differing effects on different types of academic tasks (Bull, Espy, & Wiebe, 2008; Latzman, Elkovitch, Young, & Clark, 2010).

The distinction between interference and deficit models need not necessarily make them mutually exclusive (Musch & Bröder, 1999; Naveh-Benjamin et al., 1987). Researchers have acknowledged that the impact of test anxiety can spread across the whole learning cycle (Cassady, 2004; Zeidner, 1998). As such, mechanisms of interference can affect cognitive performance during test-taking in similar ways as the experience of state anxiety, while mechanisms of deficit can impair the acquisition of information during the test preparation in similar ways as trait anxiety. Together, interference and deficit models highlight how a variety of specific cognitive, behavioural, and physiological processes may predispose or mediate the relationship between test anxiety and test performance. However, the increasingly specific examination of different components of working memory, and their influence on different tasks, inherently limits the generalisability of the findings as well as the ability for the results to be replicated.

From a practical perspective, it may be difficult for parents and teachers to readily recognise problems with students' working memory or other information processing processes. Further, identifying the role of specific mechanisms of working memory may provide little insight to mental health clinicians or school counsellors regarding how to detect, prevent, or reduce students' experience of test anxiety. It would be more clinically relevant for test anxiety to be studied in conjunction with broader emotional states, which are more readily identifiable (e.g., Pekrun et al., 2004). Hence, the present study aims to identify how general

emotional distress may be expected to influence future outcomes, in terms of both the experience of test anxiety and subsequent test performance. In examining this question, there is a need to account for the potentially confounding effects of variables that are closely related to test anxiety. The next section presents the argument that the effects of depression and anxiety need to be accounted for in order to identify the unique role of test anxiety.

1.5 Depression and Anxiety as Correlates of Test Anxiety

This section discusses three main reasons why there is substantial potential for symptoms of depression and anxiety to confound any findings relating to test anxiety. First, depressive mood disorders and anxiety disorders are the most prevalent types of mental health diagnoses, and it is anticipated that sub-clinical levels of such emotional disturbances are present in community samples. Second, the conceptual overlap between depression and anxiety disorders are widely acknowledged, and similarly these overlaps cross onto test anxiety. Third, depression and anxiety have been identified to influence academic functioning. Thus, without statistical control, any identified influence of test anxiety on test performance may simply be due to its connection with depressive or anxiety symptoms.

1.5.1 Prevalence and Comorbidity of Depression and Anxiety

Major depressive disorder has been observed to have the highest lifetime prevalence (almost 17 percent) of any single psychological disorder (Harvard School of Medicine, 2005; Sadock & Sadock, 2007), and it is the most prevalent disorder among the 16-24 age group in the United States (15.3% lifetime prevalence; Kessler & Walters, 1998) as well as New Zealand (15.1% lifetime prevalence; Oakley Browne, Wells, Scott, & McGee, 2006). Anxiety disorders, with all its different subtypes considered together, have a 23.9% lifetime prevalence for this age group in New Zealand (Wells et al., 2006).

The comorbidity between anxiety and mood disorders has been identified as the most common diagnostic overlap, with particularly common co-occurrences between Major depressive disorder and Panic disorder, Obsessive-compulsive

disorder, and Social anxiety disorder (Clark, Watson, & Reynolds, 1995; Sadock & Sadock, 2007). A national survey in New Zealand revealed that almost half (49.6%) of those with a mood disorder also had an anxiety disorder within the same 12-month period (Scott, McGee, Oakley Browne, & Wells, 2006).

The issue of comorbidity presents important implications in terms of health outcomes. Studies comparing depressed outpatients with and without a comorbid anxiety disorder show poorer prognosis for the comorbid group in terms of higher severity of psychopathology, delayed onset of therapeutic gains, and higher rates of premature termination of psychotherapy (C. Brown, Schulberg, Madonia, Shear, & Houck, 1996; Simon et al., 2004). Similar research demonstrates how the depression-anxiety comorbidity is related to poorer health outcomes such as substance use (Kauer-Sant'Anna, et al., 2007), suicide attempts (Vieta et al., 2000), and more generally, poorer psychosocial functioning and quality of life (C. Brown et al., 1996; Kauer-Sant'Anna, et al., 2007; Reich, Warshaw, Peterson, & White, 1993).

While the nomenclature in the DSM-IV-TR separated depression and anxiety as distinct disorders, there are overlaps and similarities in the way that these disorders are clinically defined and understood (Brady & Kendall, 1992; Seligman & Ollendick, 1998). In his summary of conference papers about the proposed diagnostic classifications of Major Depressive Disorder and Generalised Anxiety Disorder in the DSM-5, First (2007) summarised that these disorders share common characteristics such as their antecedents and manifest symptoms. For example, family and twin studies support a genetic basis to the neuroticism personality trait, which has been argued to underlie the comorbidity between depression and anxiety (Hettema, 2008). Early childhood experiences such as being in an adverse family environment (Moffitt et al., 2007) and having little control over the environment (Chorpita & Barlow, 1998) are also risk factors common to both types of disorders.

Anxiety and depression also share similarities in their symptom manifestation. For example, patterns of negative self-evaluation and schemas of inability to cope may be prominent features for both depression and some types of anxiety such as social anxiety and panic disorder (Otto, Powers, Stathopoulou, & Hofmann, 2008).

Behaviourally, individuals with either disorder may engage in social withdrawal and have less exposure to pleasurable activities, whether it is due to fatigue or lack of interest as in depression, or to avoid fears of social evaluation or embarrassment as in social anxiety (APA, 2000). To draw relevance to the present study, it can be seen how features of test anxiety, such as a heighten threat appraisal of failure or diminished perceived ability to cope with undesirable outcomes, substantially overlaps with features of depression and generalised anxiety.

Such similarities are reflected in the pervasive difficulty in distinguishing depression and anxiety using psychometric tools (Seligman & Ollendick, 1998; Silverman & Ollendick, 2005). For example, a meta-analysis found that widely used measures of anxiety symptoms in children and adolescents [e.g., Revised Children's Manifest Anxiety Scale (RCMAS; C. Reynolds & Richmond, 1985), State-Trait Anxiety Inventory for Children (STAI-C; Spielberger, 1973), and Child Behaviour Checklist (CBCL; Achenbach & Rescorla, 2001)] can distinguish between subtypes of anxiety disorders, but not between anxiety and affective disorders (Seligman, Ollendick, Langley, & Baldacci, 2004). An evaluation of more recently developed instruments [e.g., Screen for Child Anxiety Related Emotional Disorders (SCARED; Birmaher et al., 1997), Multidimensional Anxiety Scale for Children (MASC; March, 1997), and Spence Children's Anxiety Scale (SCAS; Spence, 1998)] with a community sample of adolescents found that they all significantly correlated around 0.70 with depression measures (Muris, Merckelbach, Ollendick, King, & Bogie, 2002).

A possible contributing factor for this is poor discriminant validity or lack of specificity to differentiate between the two disorders, such as when there is overlapping item content (Brady & Kendall, 1992; Muris et al., 2002; Stark & Laurent, 2001). Depression and anxiety may present with the same symptoms, which may be captured by the psychometric tools for assessing either disorder. For example, "I have trouble sleeping every night" is an item in the Children's Depression Inventory, and is practically identical to "It is hard for me to get to sleep at night", an item in the anxiety measure, RCMAS. To examine the effect of overlapping item content, Cole, Truglio, and Peeke (1997) isolated the overlapping items from the RCMAS and CDI and discovered that they contributed to about 12%

of the shared variance between anxiety and depression, with much of the covariance unaccounted for. Hence, the observed overlap between depression and anxiety appears to be beyond the lack of specificity in psychometric measures. Such findings raise further questions over the usefulness of viewing depression and anxiety as distinct disorders (Clark & Watson, 1991; Feldman, 1993; Seligman & Ollendick, 1998).

Some researchers conceptualise depression and anxiety as belonging to the same spectrum of affective disorder, or that they are alternative manifestations of a common underlying diathesis (Chorpita, Daleiden, Moffitt, Yim, & Umemoto, 2000; Clark & Watson, 1991; Lonigan, Carey, & Finch, 1994). Ellard, Fairholme, Boisseau, Farchione, and Barlow (2010) presented their argument for a unified therapeutic approach that focuses on the affective dysfunction that underlies the comorbidity. This follows the increasing attention being given to the role of emotion or affective states in explaining the high degree of overlap and co-occurrence of mood and anxiety disorders. An early example of this view was given by Tellegen (1985), who conceptualised two main dimensions of affective experiences: positive affect and negative affect. It was identified that negative affect was the common affective experience shared by individuals with depression or anxiety, but that the lack of positive affect was relatively specific to depressed individuals – a finding that was subsequently replicated in later studies (e.g., Dyck, Jolly, & Kramer, 1994; Watson, Clark, & Carey, 1988). Tellegen's concepts were adopted and expanded by Clark and Watson (1991) who developed the Tripartite Model of Emotion to reconcile the similarities and differences between anxiety and mood disorders.

1.5.2 Impact of Depression and Anxiety on Academic Performance

Given that there are anticipated overlaps between the features of test anxiety with depression and other anxiety subtypes, it is likely that the influence of test anxiety on test performance is confounded by the influence of these forms of emotional disturbance. This follows the extant research into how mood disturbances associated with depression and anxiety negatively correlated with academic performance. Depression during childhood and adolescence is associated with various social, behavioural, and academic problems (Garber & Horowitz, 2002;

Lewinsohn & Essau, 2002). A study of over 2000 13-17 year olds found that depressed students had lower academic achievement and greater declines in GPA (Fröjd et al., 2008). University students with depression have higher rates of academic failure (Haines, Norris, & Kashy, 1996), and were found to have missed classes, exams, and assignments significantly more often than their non-depressed counterparts (Hysenbegasi et al., 2005). Further, King et al. (1995) found that self-reported depressive symptoms were significantly associated with higher levels of test anxiety for young adolescents, which in light of extant research into the detrimental effects of test anxiety, suggests depression to have an indirect influence on test performance. Indeed, the relationship between depression and academic outcomes appears to be strongly mediated by related factors such as self-efficacy (Zychinski & Polo, 2010) and executive functioning (Owens, Stevenson, Hadwin, & Norgate, 2012).

An area of uncertainty is the extent of a temporally bidirectional relationship where depression may both be the cause and the result of poorer functioning. Indeed, researchers have highlighted that poor functioning may also be a risk factor for depression (Eley & Stevenson, 2000; Lazaratou, Dikeos, Anagnostopoulos, & Soldatos, 2010). For example, failing to achieve an academic goal has been argued to contribute to a series of emotional disturbances that later lead to depressive symptoms, such as self-defeating thoughts and feelings of discouragement (Bandura, Pastorelli, Barbaranelli, & Caprara, 1999; McDonald, 2001).

In terms of anxiety, the literature generally identifies a negative association where higher anxiety predicts poorer outcomes in terms of cognitive functioning, exam grades, and premature termination of education (Hughes, Lourea-Waddell, & Kendall, 2008; Owens et al., 2012; van Ameringen, Mancini, & Farvolden, 2003). There is also evidence that trait anxiety serve as the precursor to the experience of test anxiety for children and adolescents (Beidel et al., 1994; King et al., 1995), as well as lower cognitive functioning and academic outcomes (Fletcher et al., 1997; Owens et al., 2008).

1.6 Summary

This chapter reviewed the theoretical foundations of test anxiety and provided an outline of how changes in its theoretical conceptualisation were also coupled with developments in ways of measuring this construct. The extensive evidence on the detrimental effects of the cognitive aspects of test anxiety on test performance was presented. However, studies from the fields of educational psychology and cognitive psychology have expanded this relationship, revealing the complexity surrounding the numerous mediating mechanisms that determine how emotional experiences influence educational outcomes. From the clinical psychology perspective, this chapter presented the argument that symptoms of depression and anxiety should be taken into account when examining the connections between general emotional distress, test anxiety, and test performance.

As briefly mentioned, the Tripartite Model of Emotions offers a theoretical framework on how general emotional distress factors may account for the similarities and differences between depressive and anxiety symptoms. As such, it is highly relevant to the present study's main objective: to examine the influence of general emotional distress on test anxiety and test performance, while controlling for the confounding effects of depression and anxiety. Concurrently, this study also has the potential to examine the applicability of the model to a New Zealand sample of older teenagers. Hence, the following chapter provides an overview of the tripartite model and its empirical status.

Chapter Two The Tripartite Model of Emotion

This chapter will first review the underlying rationale for the initial development of the Tripartite Model of Emotion (Clark & Watson, 1991), which was to reconcile the similarities and differences between depression and anxiety. The specification and assumptions of the tripartite model will be discussed, followed by a review of its applicability across different populations.

2.1 Assumptions of the Tripartite Model of Emotion

Consistent with the argument that depression and anxiety have both common and distinctive features, the tripartite model of emotion proposes three higher-order constructs to explain the similarities and differences between depression and anxiety. Central to the model is the general distress construct of Negative Affect (NA), which encapsulates a range of negative mood states such as anger, loneliness, nervousness, and disgust. These unpleasant experiences are proposed to be the shared common component that is experienced by individuals with either depression or anxiety, hence contributing to the high associations between the two disorders (Watson et al., 1995).

A large survey conducted with a representative sample of 4039 youths in Australia supported the role of negative affect and the tripartite view on the underlying latent components of depression and anxiety (Tully, Zajac, & Venning, 2009). This study conducted confirmatory factor analysis of the Depression Anxiety and Stress Scale (DASS-21; Lovibond & Lovibond, 1993), which was designed to assess and discriminate depression, anxiety, and tension/stress symptoms. Consistent with the tripartite model, the best fitting model indicated that general negative affect substantially accounted for the relationship between depression and anxiety. Similar results were found across clinical (Chorpita & Daleidin, 2002; T. Brown, Chorpita, & Barlow, 1998; Joiner, Catanzaro, & Laurent, 1996) and non-clinical populations (Jacques & Mash, 2004; Laurent et al., 1999; Lonigan, Hooe, David, & Kistner, 1999).

The two other constructs of the tripartite model, Positive Affect (PA) and Physiological Hyperarousal (PH), serve to differentiate the two types of disorders. PA represents pleasurable engagement with the environment and the experience of pleasant mood states such as joy and excitement, and it is argued that the relative absence of PA is the distinct feature of depression and one that differentiates it from anxiety. PH represents somatic tension and discomfort such as pains, dizziness, and breathlessness, which is argued to be the distinctive factor specific to anxiety disorders (Joiner, Steer, et al., 1999).

On top of providing a theoretical basis for explaining the comorbidity between depression and anxiety, the tripartite model's specification of PA and NA as higher-order affective components to depression and anxiety also presents at least two major implications. Firstly, the model may be applied in the clinical setting to aid the differential diagnosis of depression, anxiety, and other psychological disorders (Watson et al., 1995). For example, consistent with the model's specifications, low levels of PA together with high levels of NA were found to be specific indicators for individuals with depression, and not anxiety disorders (Lonigan et al., 1994) or externalising disorders such as conduct disorder or ADHD (Joiner & Lonigan, 2000). Secondly, the theoretical underpinning and statistical analyses of the model's proposed relationships enable greater understanding into the extent to which NA and PA may predict future onset or progression of affective dysfunctions (T. Brown et al., 1998; Lonigan, Phillips, & Hooe, 2003). For example, Joiner and Lonigan (2000) established regression formulas to determine how combinations of NA and PA would lead to increases in anxiety and depressive symptoms in a sample of child and adolescent inpatients. They found that high NA together with low PA predicted prolonged maintenance or exacerbation of depression symptoms, while such a combination was not associated with changes in anxiety symptoms.

2.2 Measurement of the Tripartite Factors

Early studies of the tripartite model for children and adolescents were limited by the lack of specific measures for the tripartite components, and often used existing measures of anxiety and depression to define NA, PA, and PH. Two commonly used

tools were the Children's Depression Inventory (CDI; Kovacs, 1981) and the Revised Child Manifest Anxiety Scale (RCMAS; e.g., Chorpita, Albano, & Barlow, 1998; Joiner et al., 1996; Lonigan et al., 1994; Stark & Laurent, 2001). Over time, the RCMAS and CDI were noted to have a particularly strong focus on capturing the NA construct, with a relative weakness in representing the PA and PH constructs. Indeed, it was argued that the lack of strong indicators (or items) of PA and PH in studies using the CDI and RCMAS may have inappropriately and prematurely invalidated the tripartite model (e.g., Ollendick, Seligman, Goza, Byrd, & Singh, 2003). Further, there is the problematic issue of presenting tautological arguments where the measures used to capture the tripartite components, and subsequently to validate a set of proposed relationships to depression and anxiety, were themselves based on measures of depression and anxiety.

Researchers have since acknowledged the need to use measures that represent the tripartite factors in their purest affective form (Chorpita, Daleiden et al., 2000; Joiner et al., 1996; Laurent et al., 1999). One such measure is Watson, Clark, and Tellegen's (1988) *Positive and Negative Affect Schedule* (PANAS), which provides two 10-item scales for measuring levels of PA (e.g., enthusiastic, active, and alert) and NA (e.g., anger, disgust, and guilt). Initial efforts to make the PANAS better suited for the child and adolescent population were made by Joiner et al. (1996), and a final child version (PANAS-C) was developed and validated by Laurent et al. (1999). Chorpita, Daleiden et al. (2000) later developed the Affect and Arousal Scale (AFARS) with the aim of capturing the tripartite factor as 'pure' affective dimensions that did not relate specifically to anxiety or depressive symptoms.

2.3 Empirical Research on the Tripartite Model

The tripartite model was initially proposed and developed based on adult populations, and the proposed structural relationships have been validated across the adult (Beck, Benedict, & Winkler, 2003; Joiner, Steer, et al., 1999; Watson et al., 1995) and older adult age groups (Cook, Orvashcel, Simco, Hersen, & Joiner, 2004; Teachman, Siedlecki, & Magee, 2007). For example, Joiner (1996) identified the general distress factor of negative affect represented for the overlap between

depression and anxiety symptoms for a sample of undergraduate students, and that they are distinguishable by the level of positive affect and physiological hyperarousal, respectively, consistent with the three-factor tripartite model. Similar results were gathered by Watson et al. (1995), who also found that deemphasising the importance of negative affect and focusing on the distinctive factors of PA and PH helped differentiate symptoms of anxiety and depression in a clinical adult population, and supported the clinical utility of the tripartite model.

The applicability of the tripartite model to the child and adolescent population has also been studied over the past two decades, but the accumulated body of research reveals several areas of inconsistency that warrant further investigation (Anderson & Hope, 2008). Patterns have emerged regarding how specific relationships among the tripartite factors with depression and anxiety differ across age. For example, Jacques and Mash (2004) found NA to be significantly related to anxiety only for high school children (aged 15 - 19) but not primary school children (age 8 – 11). Similarly, Chorpita (2002) identified that, within the child and adolescent population from age 8 to 18, older children displayed a greater level of discrimination between depression and anxiety, producing greater fit to the model than the younger participants. This is consistent with the argument among both the clinical and psychometric literature that anxiety and depression might be expressed as a single construct in younger children, with greater differentiation occurring at later stages of development (Cole et al., 1997; Joiner et al., 1996; de Bolle, Decuyper, de Clercq, & Fruyt, 2010). Such findings caution the utility of the tripartite model in the differentiation of depression and anxiety for younger age groups.

The relationship between NA and PA is another area of inconsistency. Clark and Watson (1991) initially specified NA and PA as orthogonal dimensions – that is, they are not considered to be opposite ends of a spectrum, and individuals may experience high or low levels of one dimension irrespective of the other. For example, an individual may experience high NA (e.g., agitated) or low NA (e.g., relaxed), together with either high PA (e.g., interested) or low PA (e.g., sluggish; Clark & Watson, 1991; Tuccitto, Giacobbi, & Leite, 2009). Inherent to this argument

is the assumption of a zero-correlation between levels of NA and PA. For example, Lonigan et al. (1999) found a non-significant correlation between NA and PA (r = -0.07) for their community adolescent sample, and identified an orthogonal model (where NA and PA were not specified to correlate) provided better fit compared to an oblique or correlated model. This finding was replicated in a later study which also found the orthogonal nature of NA and PA to be stable over seven months (Lonigan et al., 2003). However, such an orthogonal relationship between NA and PA is not consistently observed. Studies with clinical child and adult populations identified how PA and NA are significantly correlated at a weak to moderate magnitude (Joiner et al., 1996; Watson, Clark, & Tellegen, 1988). Similarly, there were significant correlations between PA and NA in community samples ranging from -.15 to -.31 (Crawford & Henry, 2004; Jacques & Mash, 2004; Kiernan, Laurent, Joiner, Catanzaro, & MacLachlan, 2004)

Another aspect of contention reflects one of the tripartite model's major shortcomings in that it does not account for the differences among the subtypes of anxiety in the anxiety disorder spectrum. For example, while an elevated fear of harm is generally self-directed for individuals with specific phobias, the fear may be directed towards parents or caregivers in the case of separation anxiety, or may be towards the fear of causing harm to others in the case of obsessive-compulsive disorder (Carr, 2006). Individuals with any particular anxiety disorder tend to show idiosyncratic combinations of problems, and as a group are highly heterogeneous in terms of manifest symptoms and internal experiences (Williams, 2005). Hence, the relations of the tripartite components to anxiety may not be uniformly consistent across the anxiety disorder spectrum.

For example, Chorpita, Plummer and Moffitt (2000) found that, while generalised anxiety, panic, and obsession/compulsions were significantly related to NA, only panic and separation anxiety were significantly related to PH, suggesting that PH may not be a consistent distinguishing factor across all types of anxiety. Indeed, Mineka, Watson, and Clark (1998) proposed an integrative hierarchical model whereby the tripartite model is reconceptualised to highlight how PH is particularly specific to panic disorder over other types of anxiety (e.g., Kring, Person, & Thomas,

2007). Social anxiety or social phobia appears to be another subtype of anxiety that is particularly heterogeneous, and is more commonly observed to be negatively related to PA than other anxiety subtypes (Chorpita, Plummer et al., 2000). For example, T. Brown et al. (1998) found PA to be negatively associated with social phobia for adult outpatients. These findings were replicated with both clinical and community samples of children and adolescents (Anderson, Veed, Inderbitzen-Nolan, & Hansen, 2010; Hughes & Kendall, 2009), highlighting the potential role of low PA in the onset and experience of socially-oriented anxiety, as well as depression.

In summary, the tripartite model of emotions provides three higher-order factors of general emotional distress that have been shown to predict more specific forms of depressive and anxiety symptoms. As such, these factors of general emotional distress, such as negative affect, are highly relevant to clinical practice. The tripartite factors are also conceptual simple enough to be readily recognised by mental health clinicians, teachers, parents, or the students themselves. Thus, there are practical implications in understanding the effects of these factors on more specific outcomes in terms of test anxiety and test performance.

Chapter Three Present Study

3.1 Research Objectives

The present study is an attempt to reconcile the literature on the emotional factors of depression, anxiety, and the tripartite factors, by simultaneously examining their proposed explanatory pathways to test anxiety and test performance, as well as possible alternative or competing pathways.

This research proposes a model that incorporates the Tripartite Model of Emotion as the precursor to test anxiety and subsequently academic performance. The underlying purpose is to establish a useful heuristic framework for identifying the pathways through which emotional distress and test anxiety may impair academic performance. Structural equation modelling is a quantitative method of analysis that allows multiple linear relationships to be tested simultaneously, which serves to examine and compare the applicability of competing explanatory models.

This investigation is carried out in two main phases, each with its own set of research objectives. A cross-sectional phase of the study investigates how the tripartite factors of emotional distress may predict the experience of test anxiety. This phase also enables investigation into the applicability of the tripartite model by examining the influence of the tripartite factors on symptoms of depression and anxiety.

A subsequent prospective phase examines how the aforementioned emotional factors may have an influence on the level of performance on final external examinations. For example, specific examination of different components of test anxiety will identify whether the cognitive components of test anxiety (i.e., worry and task-irrelevant thinking) are significant predictors of lower test performance, consistent with cognitive interference theory. Also, this study aims to assess whether the higher order tripartite factors such as negative affect have mediating or direct predictive effects towards academic performance.

3.2 Hypotheses

1 The tripartite model of emotion posits that negative affect and physiological hyperarousal are the higher-order factors of emotional distress that predict anxiety problems. Hence it is hypothesised that negative affect and physiological hyperarousal will have direct influences on test anxiety, while the influence of positive affect will be non-significant.

2 Based on existing empirical evidence, it is hypothesized that test anxiety will significantly predict lower academic performance, with a small to medium effect size. Based on the cognitive interference framework, it is hypothesized that the worry components of test anxiety will significantly predict academic performance while the emotionality components will not.

3 Given the chain of relationships established by the previous two hypotheses, it is also hypothesised that negative affect and physiological hyperarousal will have indirect and negative effects on test performance, mediated by test anxiety.

Chapter Four Method

The present investigation is a quantitative study involving both cross-sectional and prospective analyses with a community sample of senior secondary school students. In the cross-sectional phase, participants completed self-report measures of test anxiety, depression, anxiety, and the tripartite factors (negative affect, positive affect, physiological hyperarousal). This occurred approximately three months before their end-of-year external examinations.

The prospective phase of the study was conducted with a sub-sample of participants who provided their consent to participate in this phase. The consenting participants' examination results were obtained from their school, and allowed investigation into how the emotional variables may have influenced test performance.

4.1 Participants

4.1.1 Cross-sectional Study

The initial sample consisted of 642 senior high school students recruited from five high schools in the mid-central region of the North Island in New Zealand. The criteria for participation were to be aged 16 or above and be in Year 12 or 13 (in other words, the final two years of their secondary education). It was assumed that the majority of these students would have completed the national qualification examinations in the year prior to this study, and thus have experiential knowledge about their level of anxiety, if any, towards academic examinations. Participation was voluntary, and their responses to the questionnaire were kept confidential.

The initial sample was made up of 246 (38.3%) male and 389 (60.6%) female students. There were 343 (53.4%) students at Year 12 and 292 (45.5%) at Year 13. The mean age of the sample was 16.86 (*SD* = 0.70), with 205 (31.9%), 316 (49.2%), 112 (17.4%), and 2 (0.3%) students aged 16, 17, 18, and 19, respectively. Seven (1.1%) participants did not provide information about their gender, age, and school year.

In terms of ethnicity, 65.0% of the sample endorsed Pākehā / NZ European as their primary ethnicity, with 10.7% endorsing primarily Māori. The Pacific Island group made up 3.7% of the sample and consisted of students from Samoa and Fiji. The East and South-East Asia group made up 6.5% of the sample and consisted of students from Thailand, Vietnam, Malaysia, China, Hong Kong, Taiwan, Japan, and South Korea. The South Asia group made up 4.2% of the sample and consisted of those from India, Maldives, Sri Lanka, Pakistan, and Iran. Participants from Europe or the United Kingdom made up 2.8% of the sample. Another 3.9% of participants endorsed multiple ethnic identities and 3.1% of participants did not provide their ethnicity.

4.1.2 Prospective Study

In order to examine how the study's variables, such as the tripartite factors, may predict academic performance, the researcher collected the results of participants' final year external examinations, which occurred approximately three months after they completed the study questionnaire. This involved a separate consent process where participants were asked by the researcher to provide their permission for their exam results to be released by the school. Thus, only a sub-sample of the initial sample of participants was involved in the prospective phase of the study. Out of the initial sample of 642 participants, 188 (29.3%) provided their consent and contributed their exam results to the study. This sample consisted of 59 (31.4%) males and 129 (68.6%) females, which was comparable to the gender composition of the initial sample (38.3% males). The average age of 16.8 years (*SD* = .722) for this sample was also comparable to the initial sample's mean age of 16.86. However, there was a drop in Māori representation from 10.7% to 6.5% due to participant attrition.

In order to identify any systematic causes for participants' non-response, the demographic data and scale scores were compared between those who provided their consent and those who did not. Independent *t*-tests revealed statistically significant group differences in levels of depression, $t(639) = 3.32$, $p < 0.01$, and negative affect, $t(630) = 2.78$, $p < 0.01$, suggesting that those who did not provide their consent to release their academic results had higher levels of these negative

emotional symptoms. However, the effect sizes of the differences were small, suggesting limited practical significance of such group differences. Thus, while there was a substantial difference in sample size between the cross-sectional and prospective phases of this study, the reduced sample was deemed largely representative of the initial sample based on demographics and symptom severity.

4.2 Measures

The questionnaire used in the present study was a compilation of instruments that measured the main components of the study's model. Measures include the *Reaction to Tests* (*RTT;* Sarason, 1984), *Reynolds Adolescent Depression Scale - 2* (*RADS-2;* W. Reynolds, 2002), *Screen for Child Anxiety Related Emotional Disorders – Revised* (*SCARED-R;* Muris, 1997), *Positive and Negative Affect Schedule for Children* (*PANAS-C;* Laurent et al., 1999), and *Affect and Arousal Scale* (*AFARS;* Chorpita, Daleiden et al., 2000).

4.2.1 Reactions to Tests (RTT)

4.2.1.1 Description

The Reactions to Tests is a measure of test anxiety based on an interference model proposed by Sarason (1984), and it represented the first shift to a multi-dimensional view of test anxiety. Consistent with his model, the RTT captures both the worry and emotionality dimensions, represented by two worry subscales (*worry* and *task-irrelevant thinking),* and two emotionality subscales (*tension* and *bodily reactions)*.

Validation studies have revealed zero-order correlations among these subscales to range from .23 to .72 (Flett, Blankstein, & Boase, 1987; Sarason, 1984), supporting the view that the four subscales measure different aspects of test anxiety. Indeed, exploratory and confirmatory factor analyses support the four-factor structure as specified by Sarason (Benson & Bandalos, 1992; Zimmer, Hocevar, Bacheloar, & Meinke, 1992).

4.2.1.2 Scoring

The RTT is a self-report measure of test anxiety which takes five minutes to complete. It contains 40 items made up of four subscales representing the four components in Sarason's model of test anxiety. The items list symptoms of test anxiety, and participants are asked to indicate whether it is typical for them to experience such symptoms on a 4-point scale (*Not Typical, Somewhat Typical, Quite Typical,* and *Very Typical*). Scores range from 40 to 160, with lower scores indicating lower levels of test anxiety.

4.2.1.3 Norms & Psychometric Properties

A community sample of undergraduate students reported a mean total score of 70.21 (*SD* = 19.85), with internal consistency at alpha = .95 (Benson & Bandalos, 1992). Sarason's (1984) validation study reported the internal consistency of the four subscales to range from .68 to .81, with a total scale reliability of .78. The study also identified the worry and task-irrelevant thinking subscales to be significantly correlated with the Cognitive Test Anxiety scale (CTA; Cassady & Johnson, 2002), demonstrating good convergent validity (Sarason, 1984). In the present study, the RTT demonstrated excellent internal consistency with Cronbach coefficient α = 0.92.

4.2.2 Reynolds Adolescent Depression Scale – II (RADS-2)

4.2.2.1 Description

The original RADS (W. Reynolds, 1986) was one of the more commonly used tools for assessing the severity of depressive symptoms in adolescents aged 11 to 20, and is frequently used in research. A normative study was carried out in New Zealand involving over 9000 high school students as part of the Youth2000 Health and Wellbeing Survey (Milfont et al., 2008). The second edition (RADS-2) contains the same items as the original RADS, but is interpreted within the four subscales identified through factor analysis (W. Reynolds, 2002). The RADS-2 manual also provides more recent normative data for a wider range of age and ethnic groups. Myers and Winters (2002) recommended the use of RADS for assessing depression in non-clinical samples due to its good psychometric properties.

The RADS-2 is used as a screen for depressive symptoms and not as a diagnostic tool for depression. It is interpreted based on empirically derived T-scores that provide an indication of the clinical severity of depression reported by the adolescent. The information provided by the RADS-2 scores can accompany clinical interviewing and assessments in consideration of a diagnosis.

4.2.2.2 Scoring

The RADS-2 is a 30-item self-report measure that takes approximately five minutes to complete. Participants are instructed to rate how often the items apply to them generally on a 4-point scale. The points are labelled *Most of the Time, Sometimes, Hardly Ever,* and *Almost Never.* The scale is made up of four subscales, being *Dysphoric Mood, Negative Self-Evaluation, Somatic Complaints,* and *Anhedonia / Negative Affect.* Scores can range from 30 to 120, with a higher value indicating higher severity of depression.

4.2.2.3 Norms and Psychometric Properties

In a New Zealand study with over 9000 participants aged between 13 and 18, the mean total score was 59.0 (*SD* = 0.6; Milfont et al., 2008). The mean score for female students of 61.6 (*SD* = 0.7) was significantly higher than for male students who had a mean score of 55.9 (*SD* = 0.5; Milfont et al., 2008). In terms of convergent validity, Milfont et al. (2008) found RADS-2 scores to be correlated positively with suicidality and negatively with happiness and general wellbeing. The original RADS items were developed based on DSM-III diagnostic criteria for depression, hence supporting its construct validity. In the present study, the RADS-2 demonstrated high internal consistency with $\alpha = .90$. Similar alphas were reported in the RADS-2 manual (W. Reynolds, 2002), with $\alpha = .93$.

4.2.3 Screen for Child Anxiety Related Emotional Disorders – Revised (SCARED-R)

4.2.3.1 Description

The original SCARED (Birmaher et al., 1997) was developed based on DSM-IV (American Psychiatric Association, 1994) diagnostic criteria for childhood anxiety disorders, which are represented in five subscales: *general anxiety, separation*

anxiety, school phobia, social anxiety, and somatic/panic. The revised version of the SCARED incorporated additional items to capture symptoms of obsessive-compulsive disorder, traumatic stress disorder, and three types of specific phobia (blood-injection type, animal type, and situational-environmental type), as represented in the DSM-IV. Together with the pre-existing items in the original SCARED, the SCARED-R measures the symptoms of the entire spectrum of childhood anxiety disorders listed in the DSM-IV (Muris, Merckelbach, Schmidt, & Mayer, 1999).

4.2.3.2 Scoring

The SCARED-R is a 66-item measure which takes around 10 minutes to complete. Items capture symptoms of anxiety disorders that can occur during childhood, and are rated on a three-point scale labelled *Never or Almost Never, Sometimes,* and *Always*. The full scale total score can range from 0 to 132, with higher values indicating higher levels of anxiety symptoms. The subscales represent the different subtypes of anxiety disorders, and are made up of differing amounts of items. These include Panic disorder (13 items), Generalised anxiety disorder (9), Obsessive compulsive disorder (9), Post-traumatic stress (4), School phobia (4), Social phobia (4), Separation anxiety (8), and three specific phobia subtypes including Blood/Injection type (7), Animal type (3), and Situational-environmental type (5).

4.2.3.3 Norms and Psychometric Properties

Two normative studies with community samples of children and adolescents reveal mean scores of 23.9 (*SD* = 13.6) and 30.4 (*SD* = 17.9) for the total SCARED-R scale (Muris, Merckelbach, Schmidt et al., 1999; Muris, Merckelbach, van Brakel, & Mayer, 1999). This is lower in comparison with normative score of 52.2 (*SD* = 21.3) for a clinical sample of youths of similar age (Muris, Dreessen, Bogels, Weckx, & van Melick, 2004). Across these studies, internal consistency for the total score was high at alpha levels above .90, with alphas above .70 for subscale scores. In the present study, the internal consistency of the total score was high at $\alpha = .94$, with subscale alphas ranging from .58 (specific phobia – situational/environmental type) to .84 (generalised anxiety disorder).

The SCARED-R has demonstrated good concurrent and convergent validity, with significant correlations with the Child Anxiety Scale (r = .88, p < .01; Muris, Merckelbach, van Brakel et al., 1999), and CBCL internalising problems (Muris et al., 2004). Compared to other more commonly used measures of anxiety (e.g., Revised Children's Manifest Anxiety Scale (RCMAS) or STAI-C), the SCARED has demonstrated good discriminant validity against those with high depression and other problems (Birmaher et al., 1997; Muris et al., 2004; Silverman & Ollendick, 2005). Having high discriminant validity is of particular importance to this study because it aims to examine the role of anxiety and depression separately.

4.2.4 Positive and Negative Affect Schedule for Children (PANAS-C) and Physiological Hyperarousal Scale for Children (PH-C)

4.2.4.1 Description

The PANAS-C (Laurent et al., 1999) is a modified version of the adult Positive and Negative Affect Scale (PANAS; Watson, Clark, & Tellegen, 1988) where the item content was changed to better suit children's reading ability and comprehension of emotions. The PANAS-C assesses both positive affect and negative affect in children and adolescents. The PH-C (Laurent et al., 2004) was later developed to measure the third tripartite component of physiological hyperarousal, and is primarily used to complement the PANAS-C. Researchers sometimes refer to the combined use of both measures as the PH-PANAS-C.

4.2.4.2 Scoring

The PH-PANAS-C is a self-report measure containing a list of 45 words, such as *interested* and *disgusted*, and asks participants to rate how much they have felt that way over the last few weeks. Ratings are based on a five point scale ranging from 1 (*Not much or not at all*) to 5 (*A lot*). The three subscales are summed separately, with higher scores indicating higher levels of the dimension. The scales consist of 12 items representing positive affect (score range: 12 – 60), 15 items representing negative affect (score range: 15 – 75), and 18 items representing physiological hyperarousal (score range: 18 – 90). The PH-C contains 18 items describing common physiological reactions of anxiety including *dry mouth, sweaty palms,* and

heart pounding. The rating format is identical to the PANAS-C, with a possible score range of 18 to 90.

4.2.4.3 Norms and Psychometric Properties

The development and validation studies conducted by Laurent et al. (1999) revealed normative scores of 26.97 (*SD* = 10.58) for NA and 43.40 (*SD* = 9.81) for PA for a non-clinical sample of school children. A clinical sample of children of comparable age reported higher levels of NA (32.10, *SD* = 15.04) and lower levels of PA (41.70, *SD* = 14.25). These scores were largely replicated in a later study of clinical youths diagnosed with an anxiety disorder, who reported mean total scores of 33.81 (*SD* = 11.08) for NA and 40.47 (*SD* = 11.39) for PA (Hughes & Kendall, 2009).

Laurent et al. (1999) reported high internal consistencies for both the positive affect and negative affect subscales, with α = .90 and .94, respectively for the validation sample, and α = .89 and .92, respectively for the replication sub-sample. Similar levels of reliability were replicated in later studies (Laurent, Joiner, & Catanzaro, 2011). Good convergent and discriminant validity were also reported, with the NA scale correlating positively with self-reports of depression (*r* = .60) and anxiety (*r* = .68), and the PA scale correlating negatively with depression (*r* = .55) and to a lesser extent with anxiety (*r* = .30). In the present study, the NA, PA, and PH scales demonstrated adequate to good internal consistency with α = .87, .90, and .84, respectively.

4.2.5 Affect and Arousal Scale (AFARS)

4.2.5.1 Description

Like the PANAS-C, the Affect and Arousal Scale assesses the three factors in the tripartite model for children and adolescents. Its items were derived from other measures such as the RCMAS, DASS, and BAS, from which some of the items had previously been identified as being relevant to the tripartite model. A particular feature of the AFARS is that its items were defined to assess affective dimensions as purely as possible, such that the items do not specifically represent the diagnostic criteria of particular anxiety or mood disorders (Chorpita & Daleidin, 2002). Initial

exploratory factor analysis identified a five-factor solution representing NA, PA, PH, anxiety, and depression, which is consistent with the tripartite model argument where anxiety and depression are distinct from the tripartite factors. The scale was further reduced by removing items that had poor factor loadings and those that were considered difficult to read and understand for the target population. A final three-factor solution was identified from the remaining 27 items, representing the three tripartite factors (Chorpita, Daleiden et al., 2000).

4.2.5.2 Scoring

The AFARS has 27 self-report items that ask participants to rate how true each statement is in relation to how they usually feel on a four-point scale, ranging from 0 (*Never True),* 1 (*Sometimes True),* 2 (*Most Times True),* to 3 (*Always True).* It consists of three subscales representing the tripartite factors: Negative Affect (8 items, score range: 0 to 24), Positive Affect (10 items, score range: 0 to 30), and Physiological Hyperarousal (9 items, score range: 0 to 27). Higher scores indicate higher levels of the relevant dimension.

4.2.5.3 Norms and Psychometric Properties

Chorpita, Daleiden et al.'s (2000) normative study was conducted with a multi-ethnic sample of 1,589 children and adolescents aged 7 to 19 years. The mean scores of the total sample for the NA, PA, and PH subscales were 8.00 (*SD* = 3.96), 21.93 (*SD* = 4.72), and 4.75 (*SD* = 3.71), respectively. The mean scores of those in the 11th and 12th grades (i.e., equivalent of the present study's sample) for the NA, PA, and PH subscales were 8.04 (*SD* = 3.65), 22.06 (*SD* = 4.22), and 3.84 (*SD* = 2.99), respectively. Internal consistency for the three subscales was moderate to high, with α = .80, .77, and .81, respectively. Similar alphas were obtained by Daleidin, Chorpita, and Lu (2000), with α = .76, .77, and .84, respectively. In the present study, the internal consistency for the NA, PA, and PH scales were α = .84, .83, and .86, respectively.

In terms of convergent validity, Daleidin et al. (2000) demonstrated that the AFARS NA scale correlated significantly to measures of worry, anxiety sensitivity, and behavioural withdrawal. The PH scale also correlated significantly to two other

measures of autonomic arousal for children. The PA scale was not significantly correlated to measures of anxiety or physiological arousal, which both indicated discriminant validity and provided support for the role of PA as proposed by the tripartite model. Similarly in a second study phase by Daleidin et al., the scores on the AFARS indicated that the relationship between the tripartite factors with depression and anxiety were largely consistent with that proposed by the tripartite model.

4.2.6 Academic Performance

The New Zealand national qualification for secondary schools is the National Certificates of Educational Achievement (NCEA). The qualification is awarded at three levels, typically representing the qualifications of years 11, 12, and 13, or the last three years of secondary education. It consists of both internal and external modes of assessment, with 'Unit Standards' representing internal course-work assignments and 'Achievement Standards' representing external examinations. Each standard is worth a prescribed amount of credits, which loosely represent the amount of time required for learning and completing the assessments for that standard (New Zealand Qualifications Authority, n.d.). Students aim to acquire enough credits to fulfil the requirements of their NCEA level.

The present study focused on the students' performance under traditional test-taking situations, and hence only collected achievement standards (henceforth *exam*) results. Exams are graded with four grade types: Not achieved, Achieved, Merit, and Excellence, each respectively representing a higher level of performance. Both educational researchers and universities assign numerical points to these grades in order to calculate a grade point average, with the most common convention being to assign Not achieved as zero, Achieved as 2, Merit as 3, and Excellence as 4 (e.g., Massey University, n.d.; Shulruf, Wang, Zhao, & Baker, 2011). The present study adopted this convention to calculate an average exam score for each participant. All available exam grades were used regardless of subject and level of the exam.

4.3 Procedure and Ethical Considerations

During the planning stages of the research, the researcher consulted with the principal, student counsellor, and teacher of one local secondary school. The consultation helped inform the design of the procedure to ensure that it was appropriate and feasible for participating schools. The school also conveyed much interest in this research topic and provided some insight into how some of their students were experiencing and dealing with test anxiety. Discussions provided clarification on the nature and format of the students' exam results, and how best to use them in the study's analysis. For example, the school staff deemed it appropriate for the exam grades of *not achieved, achieved, merit,* and *excellence* to be converted and viewed as a scaled score from zero to three, respectively.

Following approval by the Massey University Human Ethics Committee (MUHEC; Northern, Application 10/035R), the researcher contacted secondary school principals to introduce the present study, explaining the nature and purpose of the research. For each participating school, the researcher addressed an assembly of senior students and invited them to participate in the survey. During the invitation, the researcher informed the students of the voluntary and confidential nature of the survey. Students who chose to participate were given the information sheet (Appendix A) and consent form (Appendix B). Their rights as research participants were explained and participants were asked to provide their written consent by signing the consent form. The consent form also requested participants' names and email addresses, which were used to contact participants for the collection of their exam results at a later date, as well as to send them a summary of results.

Participants completed the questionnaire (Appendix C) in an assembly hall or similar venue, and the researcher was available to address any of the participants' questions during the survey. Confidentiality was protected by assigning each participant an identification number. Each individual consent form was labelled with a unique 4-digit code which corresponded to the code on the questionnaire. The code was the only identification detail on the questionnaire, which was stored separately from the consent form once they were completed by the participant.

The link between the code numbers and the participants' identities were listed in an electronic document with password protection.

The majority of surveys were conducted during the months of September and October 2010, approximately two to three months prior to the students' exams at the end of term. This period was near the end of the school curriculum, with the focus shifting towards exam preparation. After participants had been notified of their exam results in January 2011, the researcher contacted the participants through email asking for their consent to have their exam results collected for the research. A summary of preliminary results available at that point was also sent out (see Appendix D). At this time, participants were also reminded of the second phase of the study which involved the collection and analysis of examination results. The researcher requested the participants' expression of consent to obtain their examination results through their school's administrative staff. Expression of consent took the form of a reply to the researcher's email with the word *Yes*. At the end of the study, a summary of the final results (Appendix E) was sent to all participants and school principals.

Permission to use the RADS-2, PH-PANAS-C, AFARS, and the SCARED-R was granted through email correspondence with the authors, and the RTT was freely available on the author's website.

4.4 Data Analysis

Structural equation modelling (SEM) is a statistical technique involving both confirmatory and exploratory analyses of causal relationships among multiple variables, and is suited to both the testing and development of theoretical models (Byrne, 2010; Hair, Black, Babin, Anderson, & Tatham, 2006). This technique has the advantage of allowing complex intervariable relationships to be examined simultaneously from a confirmatory approach, as well as accounting for measurement errors that are inherent in self-report psychometric research (Byrne, 2010). This technique was employed in the present study to explore the validity of the hypothesized relationships among the variables in order to identify explanatory models for test anxiety and test performance using variables of emotional distress.

SEM is conducted through two main stages of analyses to examine the adequacy of models at the measurement and structural levels.

The measurement model level specifies how the study's observed variables, or questionnaire items, represent the construct they are intended to capture, as specified by the measure's theoretical underpinning (Hair et al., 2006). Evaluation of the measurement models' goodness-of-fit will reveal whether the specified factor structures adequately represent the actual data covariance matrix. Both first-order and second-order factor structures are evaluated and post-hoc analyses can be undertaken to improve the fit of the measurement models. Given the large number of observed items captured in this study, the measurement models were trimmed either through the removal or parcelling of items in order to reduce their demands on sample size, and to ensure adequate power for detecting the effects at the structural model level.

Once the measurement models are verified to be accurate in representing the latent variables, the relationships between the latent variables can then be examined through structural modelling. This stage of analysis examines whether the hypothesised models are applicable to the sample, indicated by how well the sample's data *fit* with the hypothesised set of causal relationships. Further examination of the modification indices identifies pathways of improving the proposed model and exploring the applicability of alternative, competing models (Byrne, 2010). As urged by Byrne (2010), any additional regression paths must be meaningful from a substantial and statistical standpoint. The model should be retested after each additional regression path is specified, and any further paths are specified based on the newly calculated set of modification indices.

The criteria for a parameter's statistical significance are a critical ratio greater than +/- 1.98 and a *p* value of less than .05 (Byrne, 2010). Due to the possible violation of the *independence of cases* assumption for the reduced prospective sample (explained in section 5.1.2), a stricter significance *p* level of .01 was adopted for the prospective analyses.

Following the recommendations of Hair et al. (2006), the present study employs more than one of the following types of fit indices to comprehensively evaluate the goodness-of-fit of the models: absolute fit indices, incremental fit indices, and parsimonious fit indices. Absolute fit indices provide the most basic assessment of how well the hypothesised model fits the sample data by evaluating the model independently of other possible models. Incremental fit indices compare the fit of the proposed model to a baseline model (or null model) where all observed variables are assumed to be uncorrelated and no multi-item factors are present. Parsimonious fit indices take into account the effect of model complexity in evaluating fit, and are thus useful for comparing models specified with a different number of observed variables, parameters, or constructs.

The present study employed the chi-square (χ^2) statistic, Goodness-of-Fit Index (GFI), Root Mean Square Error of Approximation (RMSEA), Comparative Fit Index (CFI), Tucker-Lewis Index (TLI), the Parsimonious Goodness-of-Fit Index (PGFI), the Parsimonious Normed Fit Index (PNFI), and the Akaike Information Criteria (AIC).

The chi-square statistic is the fundamental measure for quantifying model fit in SEM. Unlike the traditional use of chi-square in rejecting a null hypothesis, the chi-square and its associated probability value in SEM represents the likelihood of the null-hypothesis being true, in other words, that there is no significant difference between the hypothesised model and the observed data. A low chi-square value, coupled with a high probability *p* value, would indicate good fit (Byrne, 2010). However, the chi-square is a function of sample size and is also sensitive to the number of observed variables. Thus it is vulnerable to being inflated (indicating poor fit) simply due to a large sample size or when additional variables are added into the model (Hair et al., 2006).

The GFI addresses the former issue by removing sample size from its calculation, and has become a widely used indicator of fit within the present field of research (e.g., Benson, 1998; T. Brown et al., 1998). A GFI value of greater than 0.90 is indicative of acceptable fit, with 1.00 indicating perfect fit (Byrne, 2010). The RMSEA also corrects for the tendency of the chi-square statistic to reject the

specified model when there is a large sample size, and is considered the preferred indicator of fit for studies with larger samples such as the current study (Hair et al., 2006). There is a consensus that RMSEA values of greater than 0.10 indicate poor fit (Browne & Cudeck, 1993; Hair et al., 2006). Indication of good fit was specified to be less than 0.06 by Hu and Bentler (1999) and 0.05 by Browne and Cudeck (1993), who also suggested that values greater than 0.08 are indicative of mediocre fit. The χ^2/df ratio, with its inclusion of degrees of freedom, addresses the latter issue of the χ^2 statistic, namely its tendency to be inflated when there are many variables and parameters to be estimated in the model. There is a considerable range in the recommended benchmark of adequate fit for the χ^2/df ratio, such as less than 2 (Byrne, 1989), 3 (Carmines & McIver, 1981), or 5 (Wheaton, Muthén, Alwin, & Summers, 1977). Statistically, a ratio of less than 3.80 is associated with $p > .05$ (G. Brown, personal communication, February 3, 2014), indicates adequate fit.

The Comparative Fit Index (CFI) was used as the incremental fit index. It is the preferential indicator over other incremental indices such as the Normed Fit Index (NFI) because it accounts for sample size and model complexity (Bentler, 1990; Byrne, 2010). The Tucker-Lewis Index (TLI) was also employed as an incremental fit index to enable comparisons with other studies in this topic. Similar to the GFI, a CFI and TLI value greater than 0.90 is indicative of adequate fit, with 1.00 indicating perfect fit.

The Parsimony Goodness of Fit Index (PGFI), the Parsimony Normed Fit Index (PNFI), and the Akaike Information Criteria (AIC) were used to compare the goodness of fit between alternative models. These values of these indices are used for comparative purposes and are not substantively meaningful on their own. The model with a higher PGFI and PNFI values, and lower AIC value is indicative of a superior model (Byrne, 2010; Hair et al., 2006). Table 4.1 summarises the fit indices used and the adopted benchmark of good fit based on the accumulative recommendations. Hair et al. (2006) also highlighted the need to adjust these benchmarks based on different sample and model characteristics. This is particularly relevant to the present study because the structural models at the prospective phase uses a reduced sample of 188 participants.

Table 4.1
Adopted Criteria for Adequate Fit Across Cross-Sectional and Prospective Analyses

	Fit Indices							
	χ^2	χ^2/df	GFI	RMSEA	CFI	TLI	PGFI	PNFI	AIC
Cross-sectional models with *N* = 617									
	p >.05	<3.80	>.90	<.07	>.90	>.90	Higher value indicates superior model		Lower value indicates superior model
Prospective models with *N* = 188									
	p >.05	<3.80	>.92	<.08	>.92	>.92	Higher value indicates superior model		Lower value indicates superior model

*GFI: Goodness-of-fit index. RMSEA: Root mean square of approximation. CFI: Comparative fit index. TLI: Tucker-Lewis index. PGFI: Parsimony goodness-of-fit index. PNFI: Parsimony normed fit index. AIC: Akaike Information Criteria.

These are calculated using the AMOS (Arbuckle, 2009) software version 18.0.0. Maximum Likelihood Estimation (MLE) was the technique used to estimate the study's models. While alternative techniques such as Generalised Least Squares (GLS) and Asymptotically Distribution Free (ADF) may have the advantage of being less sensitive to non-normal data, MLE remains the most widely used method and has produced reliable results under many circumstances including violations of the normality assumption (Hair et al., 2006), and was appropriate for the current study's sample size.

The descriptive statistics were calculated using the *Statistical Package for the Social Sciences v16.0* (SPSS Inc., 2007). Analyses that involved structural equation modelling were conducted with the *Analysis of Moment Structures v18* (AMOS; Arbuckle, 2009) software.

Chapter Five Results – Descriptive and Preliminary Results

This chapter first describes the screening of data for accuracy, missingness, outliers, and skewness, and describes the processes undertaken to remedy problematic patterns of data. Descriptive results regarding the mean and standard deviation of scores are reported to allow gender comparisons and to establish the non-clinical nature of the current sample. Preliminary data on the intercorrelations between measures provide an indication of their psychometric properties, including the concurrent validity between the two measures of the tripartite factors (PH-PANAS-C and AFARS), and divergent validity between the measures of depression and anxiety (RADS-II and SCARED-R). Preliminary results also present basic investigation of the relationships between the study's variables and enable comparisons with existing research.

5.1 Screening

5.1.1 Missing Data

Missing data can be characterised as *missing completely at random* (MCAR), *missing at random* (MAR), and *not missing at random* (NMAR). MCAR signifies that the missing data are unpredictable and not dependent on observed variables or external circumstances. Such patterns of missingness are considered to be *ignorable* as they do not bias the dataset (Hair et al., 2006). MAR is a less stringent assumption where the pattern of missingness may be related to the study's variables, and is a common form of missing data in self-report questionnaires. Data that are MCAR or MAR can be replaced through likelihood-based imputation methods such as mean substitution, and it is also appropriate to simply remove the missing cases to allow for complete-case analyses (Tabachnick & Fidell, 2000). Missing data that are NMAR are related to a dependent variable and thus create biased patterns of missingness. This is the most problematic form of missingness and requires the missing values to be estimated through statistical modelling (Hair et al., 2006).

The extent and pattern of missing data in the current study was examined. Firstly Table 5.1 presents the frequency of missing data for each measure, and shows that the majority of items (118 out of a total pool of 211 items) had five or less missing responses (i.e., less than 1.6% of a possible total of 642 responses). The highest number of missing responses for any single item was 19 (3.0%) for "*Often I have trouble getting my breath*", followed by 15 (2.3%) missing responses for *"Often people upset me*", and 14 (1.9%) missing responses for "*I sometimes feel faint*" and "*I get upset*". All of these items were part of the AFARS measure, which had the highest frequency of missing data on average compared to the other measures.

Such patterns of missingness may reflect systemic causes of non-response that represent NMAR data, but several features of the dataset suggested that this was not the case. Specifically, the content of these items did not appear to capture particularly sensitive information that would cause participants to avoid providing a response. Further, the PH-PANAS-C, which captured the same tripartite constructs as the AFARS, had a minimal amount of missing data. These factors suggest that the pattern of missing data observed for the AFARS was not dependent on the nature of its constructs or their relationship with other variables. Thus, the relatively higher rate of missing data in the AFARS was likely due to the fact that this measure was situated last in the questionnaire, and participants were less likely to complete these items due to fatigue or lack of time. Such a pattern of missingness represented MAR and, considering the low percentage of missing responses, was not expected to cause significant biases to the data (Tabachnick & Fidell, 2000). Hence, all of the questionnaire's items were used in subsequent analyses.

Table 5.1
Missing Data by Measure

Measure	Average frequency of missing data per item	Percentage (from a possible total of 642 responses)
RTT	5.0	0.7%
RADS-II	3.9	0.6%
PH-PANAS-C	2.6	0.4%
SCARED	6.2	1.0%
AFARS	10.4	1.6%

In terms of missing data for each participant, 244 (38.0%) participants had at least one missing response in the questionnaire. As shown in Table 5.2, the majority of these participants had less than 10 missing responses across the five measures. For these participants, missing data were remedied through case mean substitution, whereby the missing values were estimated from the participant's mean of the valid data within the same subscale. This method assumes that any given data point is closely related to the scores on the other data points that capture the same construct, and is thus applicable to the self-report measures in this study (Fox-Wasylyshyn & El-Marsi, 2005). Roth, Switzer and Switzer (1999) demonstrated this method to be robust when data were missing on up to 20% of items, which again supports the use of this method for the present study.

Cases with more than 20% missing items within any subscale were often those that did not complete the particular measure or provided unrealistic patterns of responses, such as when the participant rated the same value consecutively for the whole measure, or when the ratings represented a zig-zag pattern (e.g., rating of 4, 3, 2, 1, 2, 3, 4, down the column). A total of 22 (3.4%) participants had more than 20% missing responses within any particular subscale, and were excluded listwise from further analyses.

Table 5.2
Missing Data by Participant

Number of missing responses per participant [a]	Frequency of participants	Percentage of total sample
0	398	61.9
1	140	21.8
2	54	8.4
3	22	3.4
4	5	0.8
5 → 9	7	1.1
10 → 29	5	0.8
30 → 49	4	0.6
50 → 69	5	0.8
70 → 89	1	0.2
90 → 109	1	0.2
More than 110	0	0.0
Total	642	

[a] From a total pool of 211 items.

5.1.2 Assumptions on Data Distribution

Structural equation modelling operates with a number of assumptions about the data distribution at the univariate and multivariate levels. These include the absence of univariate and multivariate outliers, normality and the absence of skewness or kurtosis, multivariate homoscedasticity, and independence of cases.

Outliers are extreme data points that typically occur when certain participants are different from the target population group (e.g., those with clinically significant distress within a non-clinical target population), or as a result of errors in data entry. The presence of outliers was identified through visual inspection of histogram and box-plot representations of the data distributions for each measure. Data points that appeared stand-alone at the extreme points in the histogram, or appeared above or below the edges of the box-plot diagram, were indicative of outliers. In order to ascertain whether any outliers had a strong influence on the data distribution, the measures' mean scores were compared with their 5% trimmed mean, where the top and bottom 5% of cases are removed. Large differences between the two means are indicative of problematic outliers. As shown in Table 5.3, the difference between the mean and 5% trimmed mean for the AFARS PH scale were potentially problematic. Inspection of the box plot for this scale revealed three participants whose ratings were approximately four standard deviations above the mean of the sample.

Table 5.3
Influence of Outliers

Measure	Mean	5% Trimmed Mean	Difference	Percentage change
RTT	85.86	85.44	0.42	0.49%
RADS	59.04	58.54	0.50	0.85%
SCARED	98.81	98.04	0.77	0.78%
PH-PANAS-C				
PA	39.33	39.60	-0.27	-0.69%
NA	27.76	27.18	0.58	2.09%
PH	30.28	29.53	0.75	2.48%
AFARS				
PA	20.86	21.17	-0.31	-1.49%
NA	8.67	8.52	0.15	1.73%
PH	4.29	3.81	0.48	11.19%

The Mahalanobis D^2 statistic was employed to detect outliers at a multivariate level, which served to identify cases with extreme or unique *combinations* of responses. The D^2 divided by the degrees of freedom (i.e., the number of scales) provided an approximate test of statistical significance, and any resulting values above a particular threshold were indicative of multivariate outliers. A threshold value of 2.5 is appropriate for small sample sizes, while a threshold value of 4.0 is more appropriate for larger samples such as the current study (Hair et al., 2006). The Mahalanobis D^2 statistic was calculated in relation to the nine scales listed in Table 5.3. The results identified five participants to be above the threshold, suggesting the presence of multivariate outliers. However, none of these five participants were also one of the three participants previously identified to be outliers at the univariate level. Following Hair et al.'s (2006) rationale, these five participants did not provide sufficiently extreme responses considered unrepresentative of the population, and thus were not removed from further analyses.

However, for the three participants who were identified as outliers at the univariate level, the magnitude of their extreme responses (approximately four standard deviations above the mean) was high enough to warrant their removal from further analyses, as the responses may indicate levels of distress above that expected from a non-clinical sample. In addition to the 22 participants removed listwise due to missing data, the resulting sample of valid cases was 617, which represented 96% of the original sample of 642 participants.

The distribution of data was screened after listwise deletion of cases with missing data (N = 617) using SPSS software. Data distributions that depart substantially from normality can cause an inflation of the $\chi 2$ statistic and the underestimation of fit indices, leading researchers to inappropriately reject or modify models that are theoretically sound (Schumaker & Lomax, 2004; West, Finch, & Curran, 1995).

Following West et al.'s (1995) convention, the data distribution is considered to be substantially non-normal at the univariate level if skewness > 2 or kurtosis > 7. As shown in Table 5.4, none of the scales reached this standard of non-normality.

Table 5.4
Univariate Normality of Distribution for All Scales and Subscales

Measure	Skewness	Kurtosis
RTT	0.35	-0.27
RADS-2	0.63	0.50
SCARED-R	0.61	-0.12
PH-PANAS-C		
Positive Affect	-0.37	-0.24
Negative Affect	0.90	0.80
Physiological Hyperarousal	1.14	1.41
AFARS		
Positive Affect	-0.77	0.41
Negative Affect	0.42	-0.18
Physiological Hyperarousal	1.48	2.23
Average Examination Grade	0.36	-0.33

Another recommended convention for identifying the extent of non-normality is to examine the visual representation of data distribution (Byrne, 2010; Hair et al., 2006). It was observed that some subscales displayed potentially problematic skewness, including the *physiological hyperarousal* scales of the AFARS and PH-PANAS-C, the *anhedonia* and *negative self-evaluation* subscales for the RADS-2, and some of the subscales of SCARED-R. As expected from using a community, non-clinical sample, most measures of psychological distress such as anxiety and negative affect were positively skewed, indicating lower levels of distress.

The assumption of multivariate homoscedasticity is met if the errors or residuals in measurement appear consistent across different values of the independent variable, and is evaluated by examining a residual scatterplot (Hair et al., 2006). For the current study, two residual plots were established to test this assumption. The first plot was based on the cross-sectional SEM analyses using the sample of 617 participants, which specified the total test anxiety score as the dependent variable in a multiple regression analysis, with the total scores of the tripartite factors, depression, and anxiety specified as the independent variables. The second plot was based on the prospective SEM analyses using the reduced sample of 188 participants, and specified exam grades as the dependent variable. The two scatterplots are provided in Appendix F, each showing a best-fit horizontal linear

regression line at zero. These provide support that the assumption of multivariate homoscedasticity has been met for the subsequent cross-sectional and prospective analyses.

The assumption of independence of cases is met when there are no discernible trends or patterns in responses among participants. The current study's design is vulnerable to violations of this assumption because the questionnaire was administered in group settings, and any common environmental condition experienced by the individuals in the group may influence their responses and inflate their correlation. The Durban-Watson statistic is used as the statistical indicator for this assumption, with values close to two indicating good independence of cases. This statistic was calculated for the two regression formula described above – one for the full sample and one for the reduced sample with data on exam grade. The Durbin-Watson statistic for the former was 1.973, which is above the critical value of 1.820 for accepting the null hypothesis at the 0.5 significance level for 200+ participants (Savin & White, 1977). This provides support for the assumption of independence for the total sample. The Durbin-Watson statistic for the reduced sample was 1.752, which is between the upper and lower critical values (namely 1.817 and 1.651) for accepting and rejecting the null hypothesis of independence, respectively. This indicates an inconclusive outcome where there is uncertainty over whether the assumption of independence has been violated (Savin & White, 1977; Spicer, 2005). Violations of this assumption likely raises Type I error and a recommended solution to account for this is to adopt a stricter level of significance to .01 (Hair et al., 2006; Stevens, 2002). This level of significance will be applied to any of the prospective analyses using the reduced sample of 188 participants.

5.2 Descriptive Results

5.2.1 Test Anxiety

Participants completed the Reactions to Tests (RTT; Sarason, 1984) as a measure of their level of anxiety towards examination situations. Table 5.5 presents the total and subscale scores as well as the intercorrelations of the RTT scales. All subscale scores were significantly correlated with the total score, with *worry* being the most strongly correlated. All intercorrelations were statistically significant, and were mostly moderate to strong in magnitude.

Table 5.5
Descriptive Data and Intercorrelations of RTT Subscales for Test Anxiety

	Mean[a] (*SD*)	Worry	Task-irrelevant thinking	Tension	Bodily sensations
Worry	23.85 (6.37)	--			
Task-irrelevant thinking	23.10 (7.25)	.42**	--		
Tension	22.83 (7.00)	.67**	.08*	--	
Bodily sensations	16.09 (4.84)	.50**	.24**	.63**	--
TOTAL	85.87 (19.06)	.87**	.61**	.78**	.75**
Normative Total[b]	70.21(19.85)				

[a]Possible score range for the subscales is 10-40, with the total score ranging from 40-160. Higher scores indicate higher levels of test anxiety. [b]Benson & Bandalos (1992), N = 318.
*p < .05. ** p <.01.

5.2.2 Depression and Anxiety

Participants completed the Reynolds Adolescent Depression Scale - II (RADS-II; W. Reynolds, 2002) as a measure of depression, and the Screen for Child Anxiety and Related Emotional Disorders – Revised (SCARED-R; Muris, 1997). The total and subscale scores for depression, as well as the intercorrelation between the RADS-II scales, are presented in Table 5.6. All subscales were significantly correlated with the total score. *Anhedonia* was weakly correlated with *dysphoric mood*, and did not significantly correlate with *somatic complaints.* All other intercorrelations were statistically significant and were mostly moderate to strong in magnitude.

Table 5.6
Descriptive Data and Intercorrelations of RADS-II Subscales for Depression

	Mean[a] (*SD*)	Anhedonia	Dysphoric mood	Negative self-evaluation	Somatic complaints
Anhedonia	12.06 (3.68)	--			
Dysphoric mood	16.80 (4.79)	.15**	--		
Negative self-evaluation	13.69 (4.75)	.35**	.69**	--	
Somatic complaints	16.42 (4.06)	.06	.67**	.58**	--
TOTAL	58.97 (13.22)	.48**	.86**	.89**	.78**
Normative total[b,c]	58.90 (0.60)				

[a]Possible score range is 7-28 for the Anhedonia and Somatic Complaints subscales, and 8-32 for the Dysphoric mood and Negative self-evaluation subscales. The total scale range is 30-120. Higher scores indicate higher levels of depression. [b]Milfont et al. (2008), *N* = 9288. [c]Standard errors reported in parentheses. *$p < .05$. **$p < .01$.

The total and subscale scores for anxiety are presented in Table 5.7. The differences in means and standard deviations between the subscales should be considered in the context of the different possible score range for each subscale. The inter-correlations between the subscale scores of the SCARED-R are presented in Table 5.8. All subscales were significantly correlated to the total scale score, with panic disorder being the most strongly correlated. The correlation between social phobia and traumatic stress did not reach statistical significance. All other inter-correlations were statistically significant, and were mostly moderate in magnitude.

Table 5.7
Descriptive Data for the SCARED-R Scales of Anxiety

Scale	Score range	Mean	*SD*
Panic disorder	0 – 26	4.36	4.04
Generalised anxiety disorder	0 – 18	7.16	4.24
Obsessive-compulsive disorder	0 – 18	5.21	3.09
Traumatic stress	0 – 8	1.50	1.96
Social phobia	0 – 8	3.06	2.16
Blood / Injection phobias	0 – 14	3.18	2.80
Animal phobias	0 – 6	.92	1.53
Situational / Environmental phobias	0 – 10	2.16	1.98
School phobia	0 – 8	2.08	1.65
Separation anxiety	0 – 16	3.08	2.85
TOTAL	0 – 132	32.73	18.04
Normative total[a]	0 – 132	30.43	17.88

[a] Muris, Merckelbach, Schmidt et al. (1999), *N* = 674. *$p < .05$. **$p < .01$.

Table 5.8
Intercorrelations of the SCARED-R Subscales for Anxiety

Scale	1	2	3	4	5	6	7	8	9	10
1) Panic disorder	--									
2) GAD	.54**	--								
3) OCD	.58**	.56**	--							
4)Traumatic stress	.60**	.43**	.50**	--						
5) Social phobia	.18**	.37**	.25**	.07	--					
6) Blood / Injection phobias	.47**	.36**	.43**	.38**	.20**	--				
7) Animal phobias	.33**	.23**	.36**	.31**	.09*	.33**	--			
8) Situational / Env. phobias	.51**	.35**	.38**	.46**	.12**	.46**	.29**	--		
9) School phobia	.48**	.37**	.30**	.32**	.20**	.34**	.19**	.35**	--	
10) Separation anxiety	.60**	.47**	.45**	.50**	.18**	.48**	.31**	.53**	.39**	--
Total scale	.83**	.76**	.75**	.68**	.40**	.67**	.47**	.65**	.57**	.75**

*$p < .05$. ** $p < .01$.

The intercorrelations among depression, anxiety, and test anxiety were found to be statistically significant. Specifically, the total SCARED-R and RADS-II scores were moderately correlated with Pearson's $r = .55$, $p < .01$, indicating a positive relationship between depression and anxiety symptoms. The total score of RTT was moderately correlated with total SCARED-R ($r = .47$, $p < .01$) and RADS-II ($r = .44$, $p < .01$), indicating that test anxiety was related to both depression and anxiety.

5.2.3 Tripartite Factors

Participants completed the Positive and Negative Affect Scale for Children (PH-PANAS-C; Laurent et al., 1999, 2004) and the Affect and Arousal Scale (AFARS; Chorpita, Daleiden et al., 2000) as the measures for positive affect, negative affect, and physiological hyperarousal. The descriptive data for the scales across the two measures are presented in Table 5.9. The intercorrelations between the AFARS and PANAS-C scales are also presented to identify the relationship among the tripartite factors, as well as provide a preliminary screen of the concurrent validity for both measures. As shown in Table 5.9, the tripartite factors were moderately and significantly correlated with their counterpart between the two measures, providing support for the concurrent validity of these measures.

Table 5.9
Descriptive Data and Intercorrelations for the Tripartite Factors

Scale (score range)	Mean (*SD*)	PH-PANAS-C			AFARS		
		PA	NA	PH	PA	NA	PH
PH-PANAS-C							
Positive Affect (12-60)	39.38 (9.39)	--					
Negative Affect (15-75)	27.69 (9.06)	-.31**	--				
Physiological Hyperarousal (15-75)	30.18 (9.42)	-.06	.51**	--			
AFARS							
Positive Affect (0-30)	20.89 (5.62)	.43**	-.07*	-.06	--		
Negative Affect (0-24)	8.68 (4.75)	-.21**	.57**	.42**	.09*	--	
Physiological Hyperarousal (0-27)	4.27 (4.58)	-.14**	.42**	.57**	-.12**	.55**	--

* $p < .05$. ** $p < .01$.

According to the tripartite model of emotion, negative affect should be the common factor between both depression and anxiety and should thus positively correlate with both. In support of this, Table 5.10 shows that RADS-II and SCARED-R scores were significantly correlated to the NA scales of both the PH-PANAS-C and AFARS, with magnitudes ranging from medium to large.

The tripartite model also posited that positive affect should be negatively correlated with depression, and should not be correlated with anxiety. As shown in Table 5.10, the PA scales of both the PH-PANAS-C and AFARS were indeed negatively correlated with RADS-II depression scores. Contrary to the tripartite model, the PA scale of the PH-PANAS-C was significantly correlated with SCARED-R, suggesting a negative relationship between positive affect and anxiety. These correlations were however small in magnitude, and appeared to apply only to certain subtypes of anxiety.

In terms of *physiological hyperarousal,* the tripartite model posited that PH should be positively correlated to anxiety but not depression. As shown in Table 5.10, the PH scales of both the PH-PANAS-C and AFARS were indeed positively correlated with SCARED-R scores. However, both PH scales were also significantly correlated with RADS-II scores, suggesting a positive relationship between depression and

physiological hyperarousal. This challenges the specificity of PH as the distinguishing factor for anxiety, and is inconsistent with the assumptions of the tripartite model.

Table 5.10
Correlations of the Tripartite Factors with Anxiety, Depression, and Test Anxiety

	PH-PANAS-C			AFARS		
	PA	NA	PH	PA	NA	PH
SCARED-R Total	**-.19****	**.54****	**.55****	**-.01**	**.68****	**.65****
Panic disorder	-.18**	.50**	.59**	-.07*	.56**	.71**
GAD	-.23**	.53**	.40**	.09*	.63**	.39**
OCD	-.13**	.46**	.37**	.04	.55**	.42**
Traumatic stress	-.11**	.36**	.37**	.02	.45**	.48**
Social phobia	-.17**	.19*	.17**	-.12**	.22**	.19**
Blood / injection phobia	-.03	.23**	.29**	.03	.39**	.41**
Animal phobia	-.04	.17**	.24**	-.02	.27**	.28**
Situation / environmental phobia	-.05	.26**	.33**	-.02	.39**	.42**
School phobia	-.25**	.35**	.38**	-.15**	.43*	.49**
Separation anxiety	-.06	.37**	.41**	.03	.48**	.47**
RADS-II Total	**-.50****	**.69****	**.44****	**-.23****	**.57****	**.42****
Anhedonia	-.50**	.25**	.14**	-.48**	.09*	.15**
Dysphoric mood	-.39**	.68**	.37**	-.03	.58**	.36**
Negative self-evaluation	-.37**	.62**	.40**	-.22**	.45**	.39**
Somatic complaints	-.28**	.49**	.40**	-.01	.55**	.36**
RTT Total	**-.13****	**.41****	**.44****	**-.03**	**.39****	**.34****
Worry	-.14**	.31**	.31**	-.01	.33**	.23**
Task-irrelevant thinking	-.03	.17**	.22**	-.07	.16**	.13**
Tension	-.13**	.37**	.33**	.05	.36**	.27**
Bodily sensations	-.11**	.42**	.51**	-.04	.36**	.45**

* $p < .05$. ** $p < .01$.

5.2.4 Academic Performance

The prospective phase of this study collected the exam results of a sub-sample of 188 participants in order to investigate how the aforementioned emotional variables may relate to academic performance. Following the convention employed by Shulruf et al. (2011) a value of 4 was given for a grade of Excellence, 3 for Merit, 2 for Achieved, and zero for Not Achieved. From this, a participant's grade point average (GPA) was calculated.

The sample obtained an average GPA of 2.00 with a standard deviation of .82. It is also noteworthy that the number of exams that each participant completed (from

which their GPA was calculated) ranged considerably from 2 to 25 exams. The implications of this variation on the interpretation of subsequent findings are discussed in Chapter 8. As shown in Table 5.11, the cognitive components of test anxiety (worry and task-irrelevant thinking) were significantly correlated with exam GPA for the total sample.

Table 5.11
Pearson's r Correlations Between Scales and Grade Point Average

Scale	*r*	Scale	*r*
Test Anxiety (RTT) Total	-.18*	Positive Affect	
Worry	-.16*	AFARS	.16*
Task-irrelevant thinking	-.39**	PH-PANAS-C	.02
Tension	.07		
Bodily Symptoms	-.05	Negative Affect	
		AFARS	.07
Depression (RADS-II) Total	-.08	PH-PANAS-C	-.09
Anxiety (SCARED-R) Total	.03	Physiological Hyperarousal	
		AFARS	-.05
		PH-PANAS-C	-.16*

* $p < .05$. ** $p < .01$.

This replicated existing research findings and provided preliminary support for the hypothesized influence of test anxiety. Scores on the other measures were mostly non-significantly correlated with exam GPA. Depending on the measure used, statistically significant relationships were identified for the tripartite factors of positive affect and physiological hyperarousal. However, the magnitudes of these correlations were weak. Some of these relationships are further examined through structural equation modelling in the following chapters.

5.2.5 Summary

Preliminary results showed that the participants' level of emotional disturbance as captured by the study's questionnaire, including symptoms of depression and anxiety, was generally comparable to other community normative samples, indicating the non-clinical nature of the sample. The exception was a notably higher level of test anxiety. There was a moderate correlation between scores of depression and anxiety, highlighting the co-morbidity and overlap between these

emotional disturbances. This significant relationship raises caution for the upcoming examination of hypothesized models that view depression and anxiety as distinct constructs.

The intercorrelations between the two measures of the tripartite factors support their concurrent validity. However, the correlation between PA and NA was significant for the PH-PANAS-C but not the AFARS, presenting inconsistencies regarding the orthogonality of these dimensions. Their relationship with anxiety and depressive symptoms offers preliminary and partial support for the tripartite model of emotions, and highlights the apparent lack of specificity of the tripartite factors as the precursors to anxiety and depression.

At a correlational level, preliminary support was found for the hypothesized role of test anxiety. However, scores of broader emotional distress did not appear to be significantly related to academic performance. These findings are further investigated in the upcoming chapters through structural equation modelling, which enables the hypothesised relationships to be examined simultaneously as well as alternative explanations to be identified.

Chapter Six Results – Measurement Modelling

As outlined in section 4.4, measurement modelling represents the first step of SEM analysis. This involves defining the study's latent constructs using the measures' underlying theoretical framework, and evaluating the adequacy of the measures in capturing the constructs they were intended to represent. This chapter describes the development and assessment of three separate measurement models: test anxiety, depression and anxiety, and the tripartite factors of emotions.

6.1 Test Anxiety

Measurement modelling first examined a hierarchical, four-factor model of test anxiety as specified by Sarason (1984). In this model (Model 1), the four latent factors were indicated by the 10 items of the corresponding subscales, and it was specified that the covariation among these factors were explained by their regression on the second-order factor of *test anxiety.* An alternative multi-dimensional model of test anxiety was then tested (Model 2) where the four first-order latent factors were specified to be intercorrelated.

Figure 6.1 shows the regression loadings for Model 1. All of the standardised regression weights for the model were significant at C.R. > +/- 1.98 and $p < 0.05$. Out of the four factors, *tension* explained the greatest proportion of the variance for test anxiety, while *task-irrelevant thinking* contributed comparatively less. The resulting fit indices are presented in Table 6.1. The RMSEA for the whole sample was .059, indicating good model fit. However, this was not accompanied by adequate levels of fit on the other absolute and incremental fit indices, and thus the model was not accepted as an adequate measurement model.

Table 6.1
Goodness-of-Fit Indices for Measurement Models of the RTT

				Fit Indices						
χ^2	*df*	χ^2/df	*p*	GFI	RMSEA	CFI	TLI	PGFI	PNFI	AIC
Model 1: Hierarchical model										
2338.760	736	3.178	.000	.817	.059	.842	.832	.734	.741	2506.760
Model 2: Multidimensional model										
2177.296	734	2.966	.000	.832	.056	.857	.848	.745	.753	2349.296

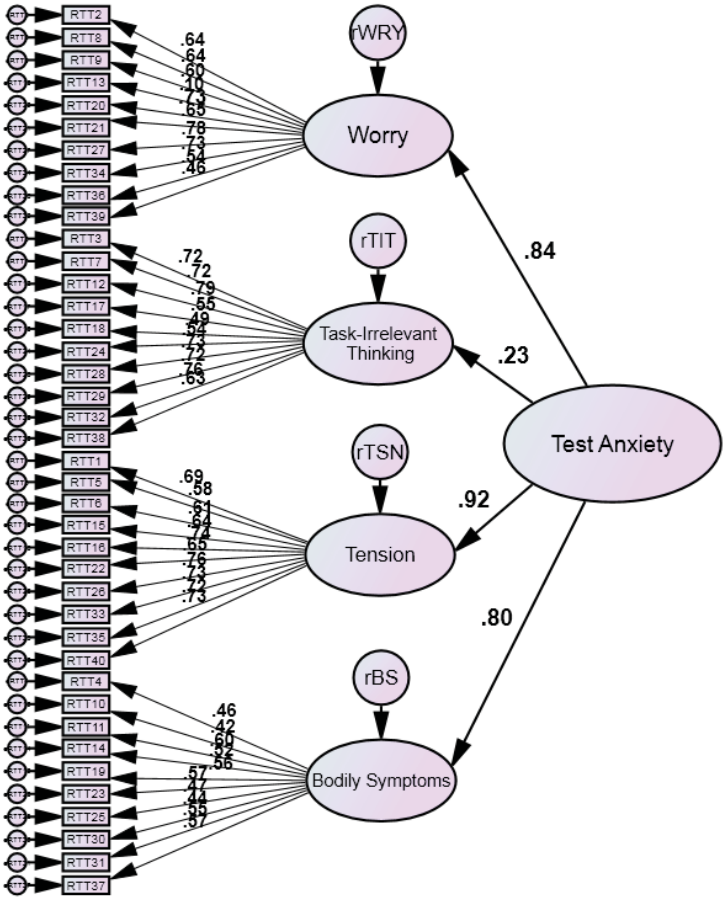


Figure 6.1. Hierarchical measurement model (model 1) of test anxiety.

An alternative first-order model was thus specified to allow the covariance between the four factors to be estimated. In other words, this represented a multi-dimensional model of test anxiety where the four first-order factors were intercorrelated (Model 2). Similar to the hierarchical measurement model, all of the observed variables' regression loadings onto their corresponding latent construct reached statistical significance. Of the factor intercorrelations, the correlation between *tension* and *task-irrelevant thinking* was not statistically significant. As indicated in Table 6.1, the model did not demonstrate adequate fit across the fit indices. Section 6.4 later describes the rationale and processes through which a trimmed measurement model of test anxiety was established.

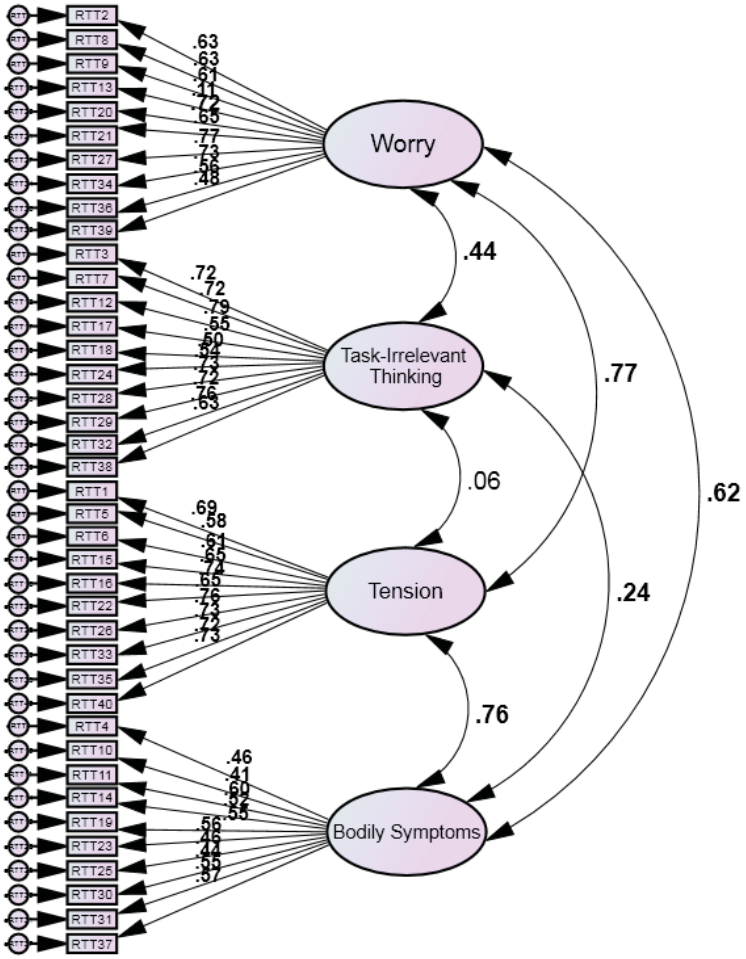


Figure 6.2. Multi-dimensional measurement model (model 2) of test anxiety.

6.2 Depression and Anxiety

Based on the literature on the overlap between depressive and anxiety symptoms, as well as the moderate correlation identified in the present study between the total RADS-II and SCARED-R scores, the constructs of depression and anxiety are anticipated to be closely related. Thus, attempts were made to establish a combined measurement model of both depression and anxiety, specified as separate but correlated latent variables. Prior to specifying this combined measurement model, the measurement models for the RADS-II and the SCARED-R were first evaluated separately in order establish whether each individual measure was adequate in capturing the constructs of depression and anxiety, respectively. Testing these separately also served to avoid having an excessive number of parameters to be estimated in the measurement model.

A hierarchical measurement model for the RADS-II was first tested by establishing *depression* as the second-order latent variable that was represented the four factors of *anhedonia, dysphoric mood, negative self-evaluation,* and *somatic complaints.* Each factor was represented by their corresponding subscale items (Model 1 / Figure 6.3). All of the observed variables significantly regressed onto their corresponding factor, and all of the factor regressions were statistically significant. Examination of modification indices did not reveal indications of misspecification other than some correlated errors between particular items. For example, there was a particularly high correlated error between item 19 (*I feel I am bad*) and item 20 (*I feel I am no good*), which appear to be due to the similarity in item content. As shown in Table 6.2, while the RMSEA value was within the range of acceptable fit, the other absolute and incremental fit indices indicated poor fit.

An alternative model specified a first-order factor structure with four intercorrelated latent variables, which represented a multi-dimensional view of depression (Model 2 / Figure 6.4). In this model, *anhedonia* did not significantly correlate with *somatic complaints* ($r = 0.71$, $p = .174$). Both models 1 and 2 suggest that *anhedonia* is relatively independent from the construct of depression compared to the other aspects. As all items in the anhedonia subscale were reverse scored, the unexpected non-significance of its relationship to the other subscales may indicate a response set bias. Table 6.2 indicate this model to be poor fitting across most of the fit indices.

Table 6.2
Goodness-of-Fit Indices for Measurement Models of the RADS-II

				Fit Indices						
χ^2	*df*	χ^2/df	*p*	GFI	RMSEA	CFI	TLI	PGFI	PNFI	AIC
Model 1: Hierarchical model										
1370.076	401	3.417	.000	.858	.063	.840	.827	.740	.728	1498.076
Model 2: Multidimensional model										
1282.523	399	3.214	.000	.868	.060	.855	.841	.745	.737	1414.523

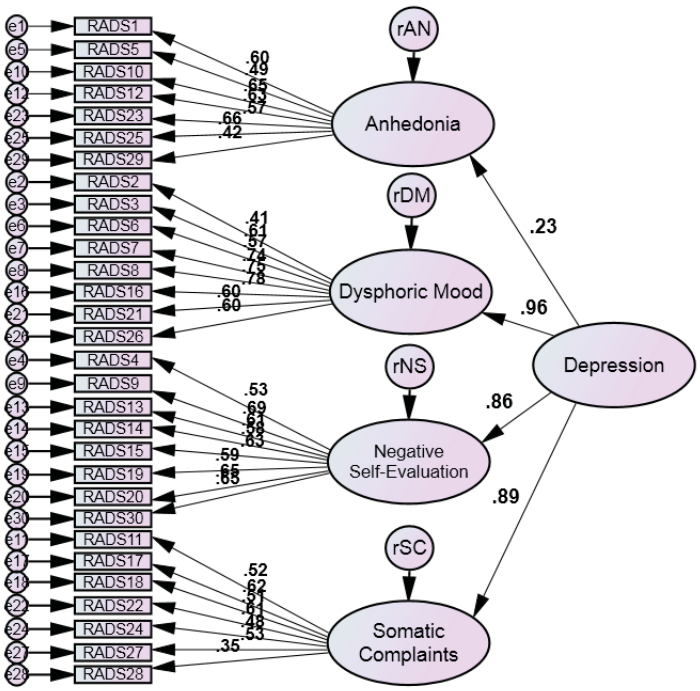


Figure 6.3. Hierarchical measurement model (model 1) of depression.

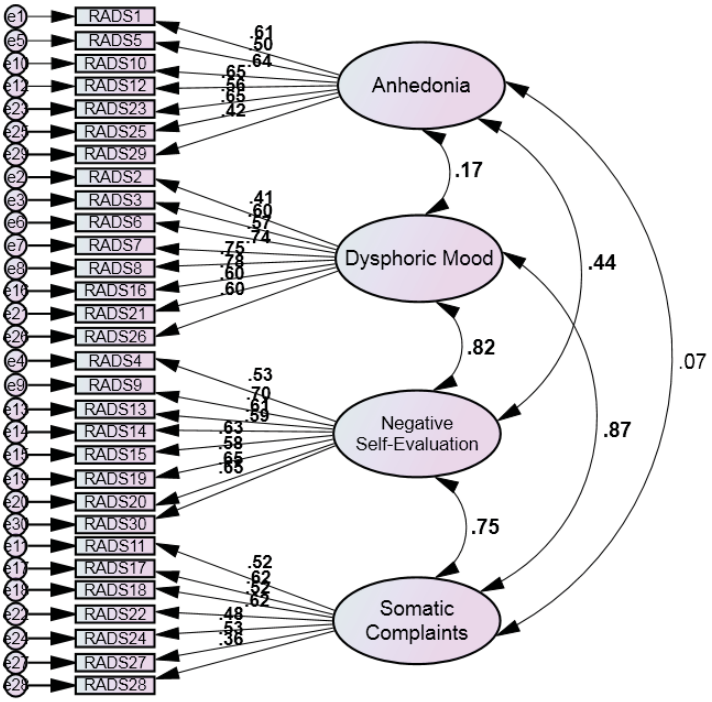


Figure 6.4. Multi-dimensional measurement model (model 2) of depression.

The adequacy of the SCARED-R in measuring anxiety was evaluated. To improve its theoretical relevance to the study, only five subtypes of anxiety were selected from the SCARED-R scale to make up a reduced measurement model of anxiety. The

subtypes were chosen based on their theoretical significance in the current literature as well as to enable comparisons with other studies (e.g., Chorpita, 2002). From this, the selected subscales were panic disorder, generalised anxiety, obsessive-compulsive disorder, social phobia, and school phobia. Both hierarchical and multi-dimensional measurement models of anxiety were examined.

A hierarchical model of anxiety was first tested, whereby the five subtypes of anxiety captured by the SCARED-R were specified as latent factors, each represented by its items, and each regressing onto a higher-order construct of anxiety (Model 1 / Figure 6.5). All parameter estimates reached statistical significance, with *panic disorder* explaining the greatest proportion of variance for anxiety. Comparatively, *social phobia* had the weakest contribution. As shown in Table 6.3, this model did not produce adequate fit consistently across the indicators. A multi-dimensional measurement model of the selected subtypes of anxiety was also specified and tested (Model 2 / Figure 6.6). All parameter estimates including item loadings and factor intercorrelations reached statistical significance. However, as shown in Table 6.3, this model was did not produce adequate fit consistently across the fit indices.

Table 6.3
Goodness-of-Fit Indices for Measurement Models of the SCARED-R

				Fit Indices						
χ^2	*df*	χ^2/df	*p*	GFI	RMSEA	CFI	TLI	PGFI	PNFI	AIC
Model 1: Hierarchical model										
1677.076	555	3.022	.000	.858	.057	.815	.802	.756	.698	1827.076
Model 2: Multidimensional model										
1605.664	550	2.919	.000	.863	.056	.826	.812	.754	.702	1765.664

In the context of these poor fit models, a combined measurement model for depression and anxiety was not established at this stage. Section 6.4 describes the process of trimming the RADS-II and SCARED-R scales in order to address the inadequate measurement models for both depression and anxiety.

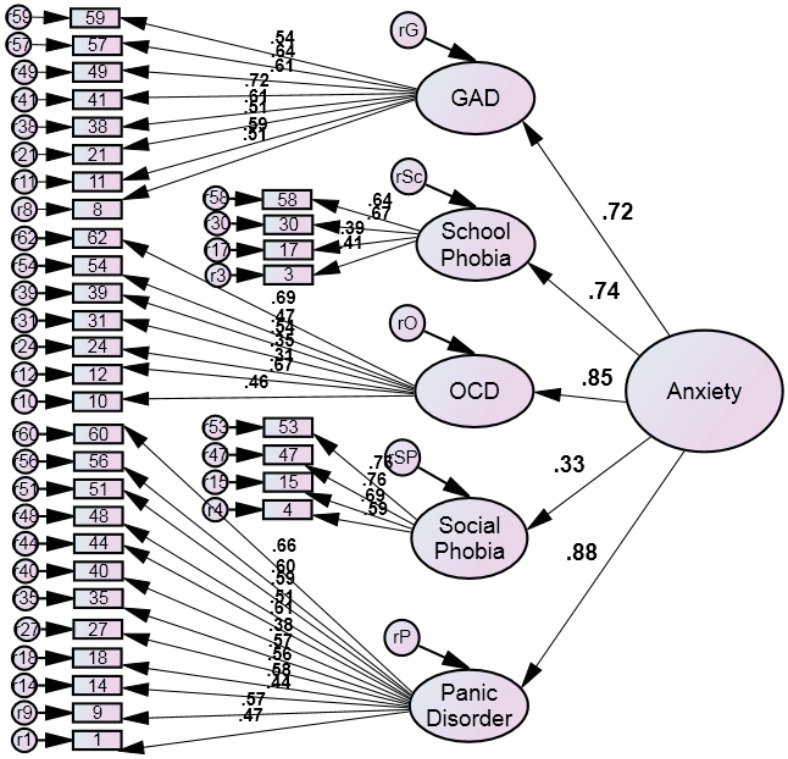


Figure 6.5. Hierarchical measurement model (model 1) of anxiety.

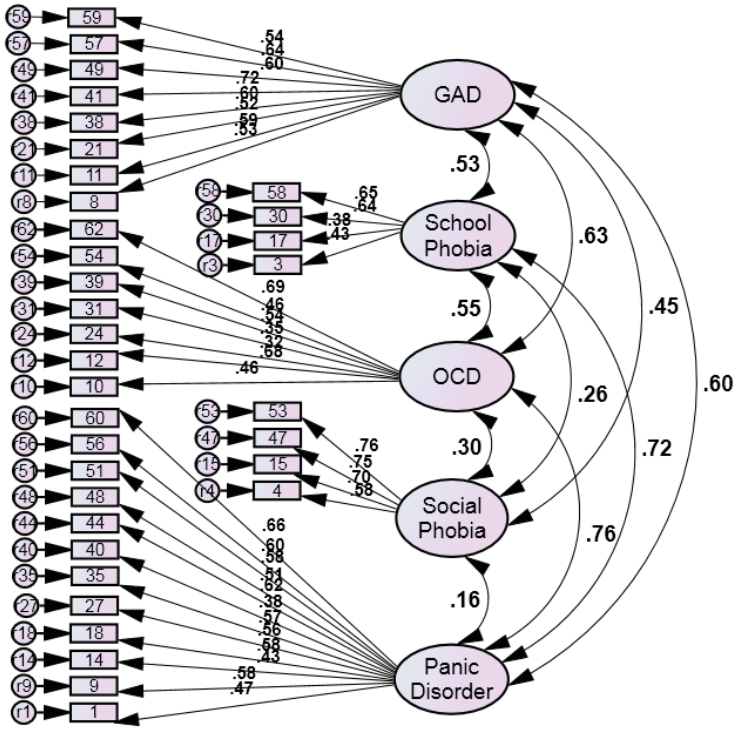


Figure 6.6. Multi-dimensional measurement model (model 2) of anxiety.

6.3 Tripartite Factors

The measurement models for the tripartite model of emotion placed Positive affect (PA), Negative affect (NA), and Physiological Hyperarousal (PH) as the inter-correlated latent factors. Measurement models were estimated separately for PH-PANAS-C and AFARS in order to examine the internal structure of each of these measures and how they represented the three tripartite components.

The model was first tested for the PH-PANAS-C. As shown in Figure 6.7, all item loadings to their corresponding factors were significant at $p < .05$. The factor inter-correlation between PA and NA reached statistical significance, which was inconsistent with the assumption of orthogonality between these dimensions (e.g., Tuccitto et al., 2010). The other factor inter-correlations also reached statistical significance, although the magnitude of the relationship between PA and PH was notably weak. As presented in Table 6.4, this model did not demonstrate adequate fit consistently across the fit indices.

The model was then tested using the AFARS. As shown in Figure 6.8, all item loadings to their corresponding factors were significant at $p < .05$. The correlation between PA and NA did not reach statistical significance, suggesting that the two factors had an orthogonal relationship. This was inconsistent with the significant relationship identified by the PANAS-C, highlighting possible differences in the way these measures represent the constructs. The other factor inter-correlations reached statistical significance, although the magnitude of the relationship between PA and PH was notably weak. As presented in Table 6.4, this model produced an adequate RMSEA value, but was slightly under the benchmark for adequate fit across the other absolute and incremental fit indices.

Table 6.4
Goodness-of-Fit Indices for Measurement Models for the Tripartite Factors

				Fit Indices						
χ^2	*df*	χ^2/df	*p*	GFI	RMSEA	CFI	TLI	PGFI	PNFI	AIC
PH-PANAS-C										
3171.445	942	3.367	.000	.798	.062	.772	.761	.726	.672	3357.445
AFARS										
957.537	321	2.983	.000	.896	.057	.886	.875	.761	.767	1071.537

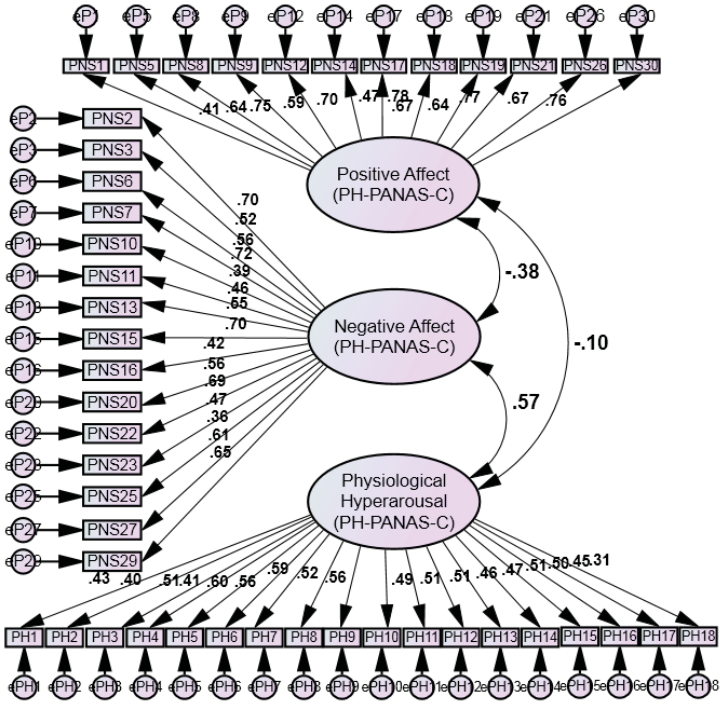


Figure 6.7. Measurement model of the tripartite factors using the PH-PANAS-C.

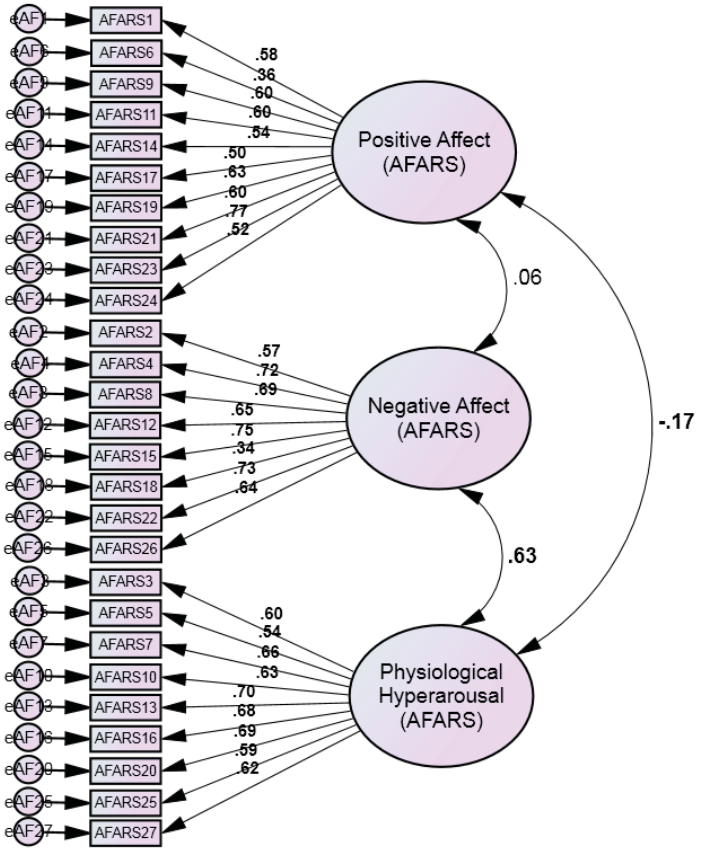


Figure 6.8. Measurement model of the tripartite factors using the AFARS.

6.4 Trimmed Measurement Models

So far in this chapter, the measurement models for each of the measures used in this study were tested for their fit to the sample data. While all of the measurement models' RMSEA values were less than .07, none of these were accompanied by CFI, GFI, or TLI values that reached the criteria of good fit. This indicates problems in the measurement models where the proposed factor loadings do not fit with the actual relationships in the data. The findings suggest that it will be inappropriate to use these measurement models to represent the latent constructs in subsequent structural models. Further, the structural model in the prospective phase involves all of the latent variables listed above, giving a large number of observed variables that will not be compensated by the reduced sample size of 188 participants. To resolve both of these issues, trimmed versions of each measurement model were tested, either through using a reduced number of manifest variables or through parcelling of item sets. These methods effectively reduce the number of parameters to be estimated and thus reduce sample demands and the effects of measurement errors. The relative cost is that the psychometric qualities of the original measures no longer apply to the new measurement models. Also, the act of removing manifest variables will effectively reduce the breadth of content captured by the model. As detailed in the upcoming sections, the selection of items and their contents are listed in Appendix G.

6.4.1 Trimmed Measurement Model for Test Anxiety

A trimmed measurement model of test anxiety was established using the four highest loading items for each of the four factors for the RTT. The content of the resulting items are provided in Appendix G, and appears to adequately retain both the breadth of content and the face validity of the original measure. Figure 6.9 presents the second-order hierarchical measurement model (model 1) for the trimmed scale. All of the manifest variables significantly regressed onto their corresponding factor at C.R. > +/- 1.98 and $p < .05$. The regression path of *task-irrelevant thinking* was not statistically significant. The goodness-of-fit indices for this model is provided in Table 6.6, and indicate adequate fit only on the GFI and CFI, with marginal or poor fit across the $\chi 2/df$ ratio, RMSEA, and the TLI.

Figure 6.10 presents the first-order multidimensional measurement model (model 2) for the reduced scale. All of the manifest variables significantly regressed onto their corresponding factor. Of the factor intercorrelations, *task-irrelevant thinking* was not significantly correlated with *tension* or *bodily symptoms,* while all other correlations were significant. Both models 1 and 2 strongly suggest that *task-irrelevant thinking* is relatively independent from the construct of test anxiety compared to the other aspects of test specific worry or emotionality (e.g., tension and bodily symptoms). The resulting fit indices for model 2 are provided in Table 6.5, demonstrating good fit across the absolute and incremental fit indices. Comparisons of the parsimonious fit indices also reveal model 2 to have superior fit.

The same model was tested using the reduced sample of 188 participants in order to clarify whether the measurement model remains robust for the prospective phase of analyses. As shown in Table 6.5, model 2 demonstrated good levels of fit across the indices for the reduced sample. Hence, the trimmed multidimensional measurement model of test anxiety, as depicted in figure 6.10, is used at the structural level for the cross-sectional and prospective model testing.

Table 6.5
Goodness-of-Fit Indices for Trimmed Measurement Models of the RTT

				Fit Indices						
χ^2	*df*	χ^2/df	*p*	GFI	RMSEA	CFI	TLI	PGFI	PNFI	AIC
Model 1: Trimmed hierarchical model (*N* = 617)										
418.633	100	4.186	.000	.914	.072	.908	.889	.672	.735	490.633
Model 2: Trimmed multidimensional model (*N* = 617)										
306.748	98	3.130	.000	.937	.059	.939	.926	.675	.746	382.748
Model 1: Trimmed hierarchical model (*N* = 188)										
154.873	100	1.549	.000	.896	.054	.949	.938	.659	.725	226.873
Model 2: Trimmed multidimensional model (*N* = 188)										
141.223	98	1.441	.003	.909	.049	.960	.951	.655	.720	217.223

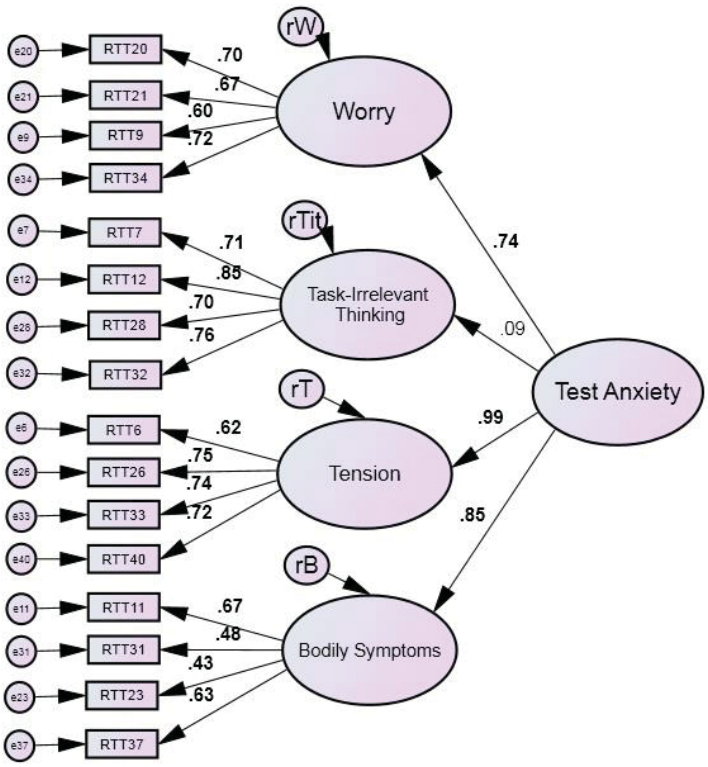


Figure 6.9. Trimmed hierarchical measurement model of test anxiety.

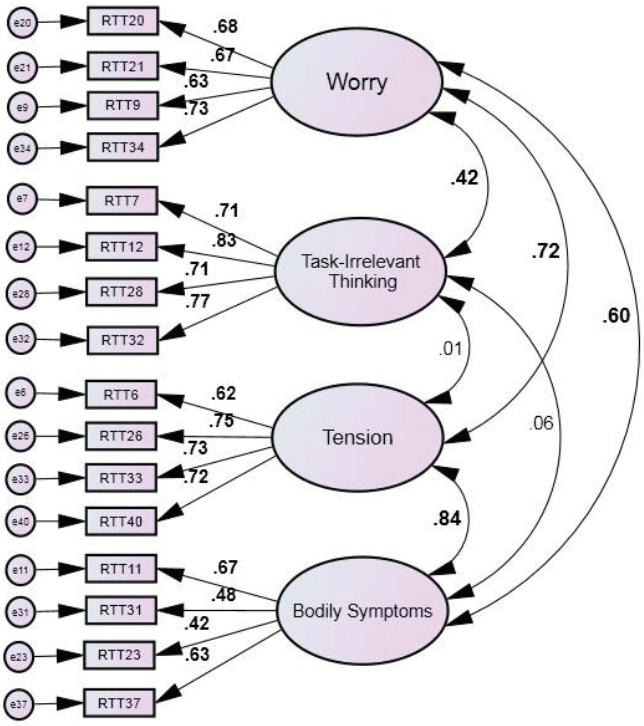


Figure 6.10. Trimmed multidimensional measurement model of test anxiety.

6.4.2 Trimmed Measurement Models for Depression and Anxiety

Attempts were made to establish a measurement model for both depression and anxiety in order for them to be included in the structural model. A trimmed measurement model for depression was established using four items for each of the four factors for the RADS-II. The highest four loading items were used to represent *anhedonia, dysphoric mood, and negative self-evaluation.* As for *somatic complaints,* the four highest loading items did not include some of the core somatic features of depression, such as "I feel tired" and "I have trouble sleeping". These items were selected over the other items with higher loadings. The content of the resulting items are provided in Appendix G, and appears to adequately retain both the breadth of content and the face validity of the original measure. Figure 6.11 presents the second-order hierarchical measurement model (model 1) for the trimmed scale. All of the manifest variables significantly regressed onto their corresponding factor at C.R. > +/- 1.98 and $p < .05$. All regression paths for the four factors were also statistically significant. The goodness-of-fit indices for this model are provided in Table 6.6, and indicate adequate fit across the fit indices.

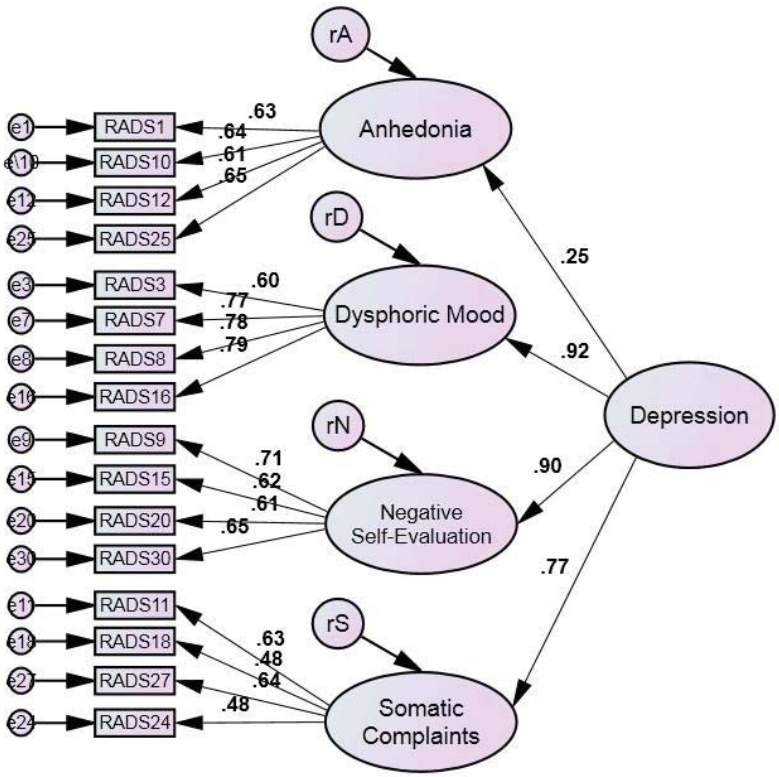


Figure 6.11. Trimmed hierarchical measurement model of depression.

Figure 6.12 presents the first-order multidimensional measurement model (model 2) for the trimmed scale. All of the manifest variables significantly regressed onto their corresponding factor. Of the factor intercorrelations, *anhendonia* was not significantly correlated with *somatic complaints,* while all other correlations were significant. The resulting fit indices for model 2 are provided in Table 6.6, demonstrating good fit across the absolute and incremental fit indices.

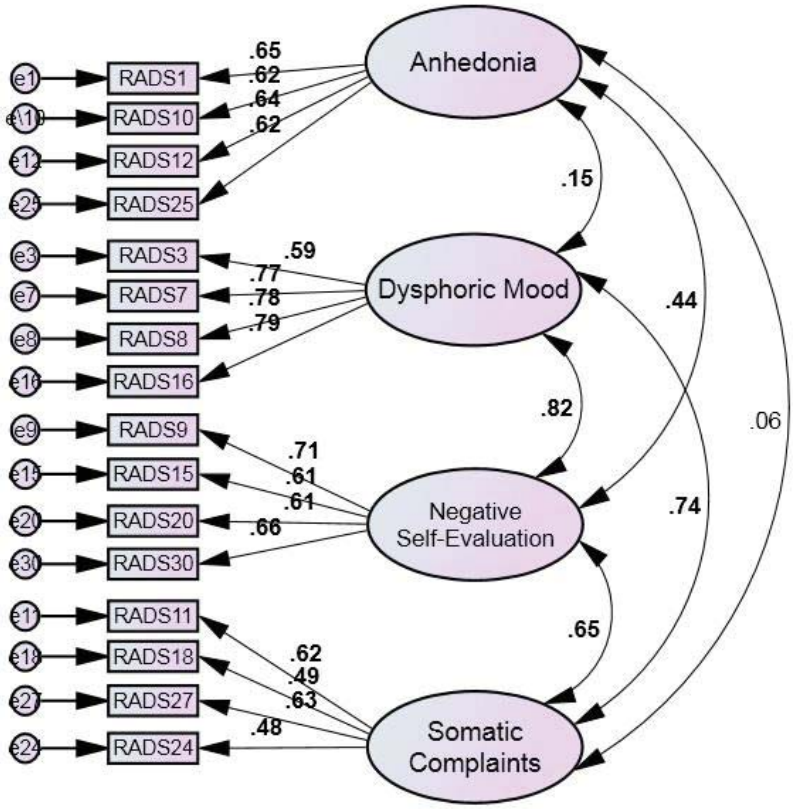


Figure 6.12. Trimmed multidimensional measurement model of depression.

Both models 1 and 2 were tested using the reduced sample of 188 participants in order to clarify whether the measurement models would remain robust for the prospective phase of analyses. There were no changes to the statistical significance of the factor regression loadings between the full and the reduced samples, and as shown in Table 6.6, both models demonstrated adequate levels of fit across the indices. The multidimensional model (model 2) was of superior fit, which is typically expected due to the greater number of freely estimated parameters.

Table 6.6
Goodness-of-Fit Indices for Trimmed Measurement Models of the RADS-II

				Fit Indices						
χ^2	*df*	χ^2/df	*p*	GFI	RMSEA	CFI	TLI	PGFI	PNFI	AIC
Model 1: Trimmed hierarchical model (*N* = 617)										
353.834	100	3.538	.000	.931	.064	.916	.900	.684	.740	425.834
Model 2: Trimmed multidimensional model (*N* = 617)										
297.812	98	3.039	.000	.942	.058	.934	.919	.679	.740	373.812
Model 1: Trimmed hierarchical model (*N* = 188)										
182.657	100	1.827	.000	.908	.066	.927	.912	.657	.711	254.657
Model 2: Trimmed multidimensional model (*N* = 188)										
166.936	98	1.703	.000	.914	.061	.939	.925	.651	.707	242.936

For pragmatic purposes, the hierarchical measurement model is used to represent depression in the structural model. This better serves the purpose of including depression in the structural model, which is to control for the influence of the tripartite factors on depression when examining their influence on test anxiety, rather than to examine the influence of the tripartite factors across the different dimensions of depression. To further reduce the number of parameters to be estimated in the structural model, item parcelling was carried out by summing the four manifest items for each of the factors to create a single manifest indicator to represent that factor. Item parcelling is considered appropriate only when the underlying factor structure of the items is clear, and that the items from which the parcel is created are unidimensional (Little, Cunningham, Shahar, & Widaman, 2002). Since the items were based on a pre-established and validated measure, it was cautiously assumed that the factor structure behind the sets of items was robust. To clarify this, the internal consistencies for each set of items were calculated, revealing good to excellent reliability coefficients for *anhedonia* (Cronbach's alpha = .723), *dysphoric mood* (.819), *negative self-evaluation* (.743), and *somatic complaints* (.710). Exploratory factor analysis was also carried out to see whether item sets represented a single, rather than multiple, factor. The resulting eigenvalues for the four items sets were 2.199, 2.611, 2.261, and 2.185, respectively, and none of the item sets produced a 2-factor solution with eigenvalues greater than 1. Both the alphas and eigenvalues indicated the

unidimensional nature of the four sets of items, and provided support the items sets to be parcelled. The resulting parcelled model is presented in figure 6.15 as a combined measurement model of depression and anxiety.

A trimmed measurement model of anxiety was established using the four highest loading items for each of the five subscale factors selected from the SCARED-R. Figure 6.13 presents the second-order hierarchical measurement model (model 1) for the trimmed scale. All of the manifest variables significantly regressed onto their corresponding factor at C.R. > +/- 1.98 and $p < .05$. All regression paths for the five factors were also statistically significant. The goodness-of-fit indices for this model are provided in Table 6.7, indicating marginal fit for the χ2/df ratio, and poor fit for the CFI and TLI.

Figure 6.14 presents the first-order multidimensional measurement model (model 2) for the trimmed scale. All of the manifest variables significantly regressed onto their corresponding factor. Of the factor intercorrelations, *social phobia* was not significantly correlated with *panic disorder,* while all other correlations were significant. The resulting fit indices for model 2 are provided in Table 6.7, indicating marginal fit for the χ2/df ratio, and poor fit for the CFI and TLI. Both models 1 and 2 were tested using the reduced sample of 188 participants to clarify whether the measurement models would fit the data for this sample. As shown in Table 6.7, both models demonstrated poorer levels of fit across the indices.

Table 6.7
Goodness-of-Fit Indices for Trimmed Measurement Models of the SCARED-R

	Fit Indices									
	χ^2	*df*	χ^2/df	*p*	GFI	RMSEA	CFI	TLI	PGFI	PNFI	AIC
Model 1: Trimmed hierarchical model (*N* = 617)											
	604.770	165	3.665	.000	.910	.066	.867	.847	.715	.719	694.770
Model 2: Trimmed multidimensional model (*N* = 617)											
	553.909	160	3.462	.000	.917	.063	.881	.859	.698	.709	653.909
Model 1: Trimmed hierarchical model (*N* = 188)											
	317.654	165	1.925	.000	.859	.070	.860	.839	.675	.653	407.654
Model 2: Trimmed multidimensional model (*N* = 188)											
	302.817	160	1.893	.000	.865	.069	.869	.845	.659	.643	402.817

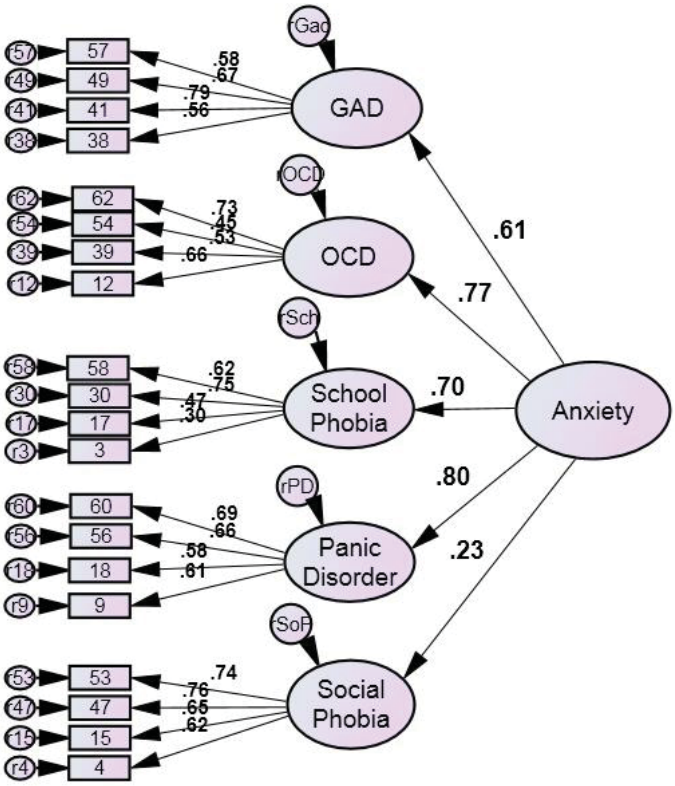


Figure 6.13. Trimmed hierarchical measurement model of anxiety.

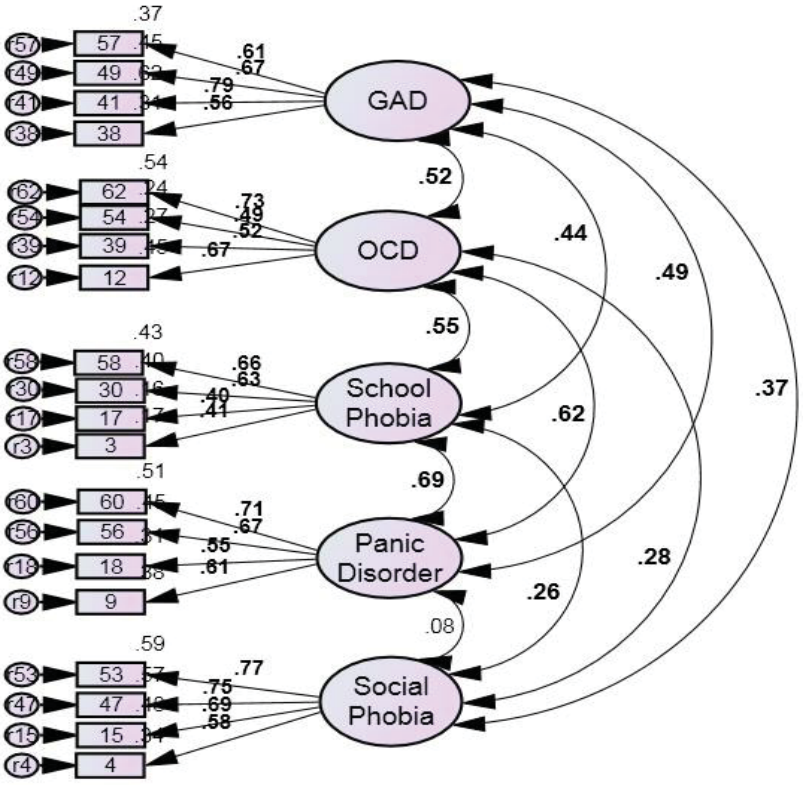


Figure 6.14. Trimmed multidimensional measurement model of anxiety.

Despite attempts to improve model fit by reducing the number of freely estimated parameters and the resulting sample demands, the current study could not establish an adequate measurement model of anxiety. This perhaps reflects the nature of the SCARED-R measure, which captures separate types of anxiety disorders rather than different dimensions of a general construct of anxiety. In a structural model, it may be more appropriate for each subtype of anxiety to be presented as separate individual latent variables with their own regression paths from the tripartite factors. However, examining the differential influence of the tripartite factors across different types of anxiety is beyond the scope of the present research objectives.

Hence, in establishing a combined measurement model for depression and anxiety, only one subtype of anxiety was selected, namely *generalised anxiety*, in conjunction with the parcelled measurement model of depression previously established. Figure 6.15 presents the resulting combined measurement model for depression and generalised anxiety (Model 1). All item loadings and the factor correlation reached statistical significance. However, as shown in Table 6.8, this model did not produce adequate fit for the χ^2/df ratio, RMSEA, and the TLI.

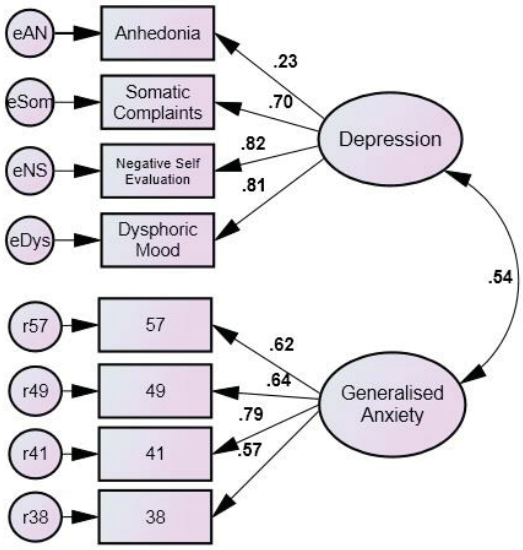


Figure 6.15. Initial measurement model (model 1) of depression and anxiety.

Review of the modification indices revealed a potential influential correlated error between a*nhedonia* and *negative self-evaluation.* Due to the theoretical similarly between these aspects, the correlated error was specified and the resulting model is presented in figure 6.16.

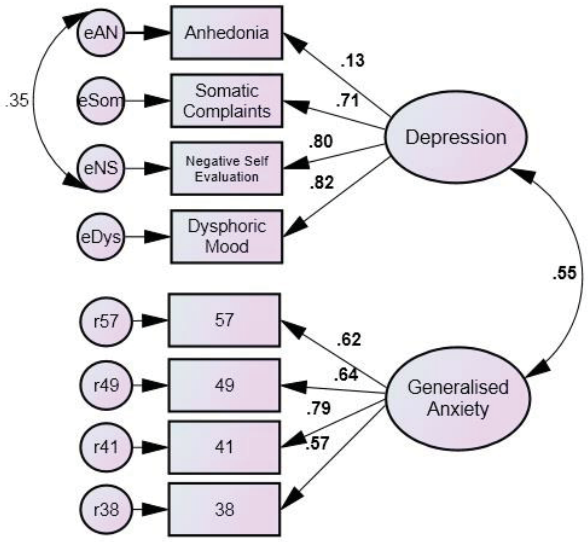


Figure 6.16. Alternative measurement model (model 2) of depression and anxiety.

As shown in Table 6.8, model 2 produced a higher level of fit across the absolute and incremental fit indices. Apart from the χ^2/df ratio, the fit indices indicated model 2 to be an adequate measurement model. There were inconsistent findings regarding the parsimonious fit indices and the AIC, where the PGFI and PNFI suggested model 1 as superior while the AIC would support model 2. Given that overall model 2 demonstrated adequate fit across most indices, it was considered the superior measurement model. Review of the modification indices suggested a correlated error between item 57 ("I worry about how well I do things") and item 38 ("I worry about things working out for me"). This additional path was not specified because the theoretical justification for doing so was not deemed strong enough to outweigh the cost of artificially inflating the fit of the model. Hence, model 2 was used as the final measurement model for depression and generalised anxiety. As explained in the subsequent chapter, depression and generalised anxiety were not included in the prospective analyses, and thus this measurement model was not evaluated for the prospective sample.

Table 6.8
Goodness-of-Fit Indices for Measurement Models of Depression and Generalised Anxiety

				Fit Indices						
χ^2	*df*	χ^2/df	*p*	GFI	RMSEA	CFI	TLI	PGFI	PNFI	AIC
Model 1: Initial Model (*N* = 617)										
135.832	19	7.149	.000	.945	.100	.921	.884	.499	.618	169.832
Model 2: Final Model (*N* = 617)										
83.191	18	4.622	.000	.965	.077	.956	.932	.483	.607	119.191

6.4.3 Trimmed Measurement Model for the Tripartite Factors

Trimmed measurement models for the tripartite factors were established for both the PANAS-C and AFARS, using the four highest loading items for each factor. Figure 6.17 the trimmed measurement model for the PANAS-C. All of the manifest variables significantly regressed onto their corresponding factor, and all of the factor intercorrelations were statistically significant at C.R. > +/- 1.98 and $p < .05$. As presented in Table 6.9, the fit indices for this model indicate a good fitting model across the fit indices. Evaluation of this model using the reduced sample of 188 participants also yielded an adequate fitting model.

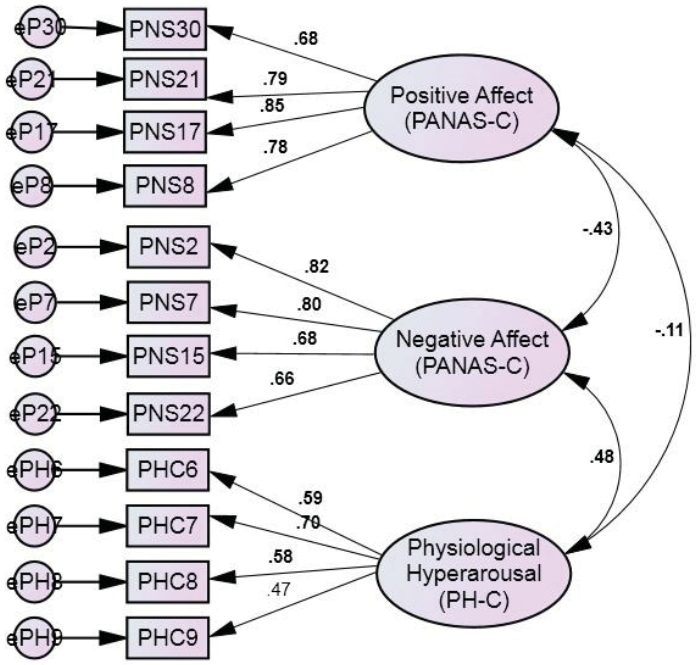


Figure 6.17. Trimmed measurement model of the tripartite factors with PANAS-C.

Figure 6.18 presents the trimmed measurement model for the AFARS. All of the manifest variables significantly regressed onto their corresponding factor. Of the factor intercorrelations, the correlation between *positive affect* and *negative affect* did not reach statistical significance, while *physiological hyperarousal* was positively correlated with *negative affect,* and negatively correlated with *positive affect*.

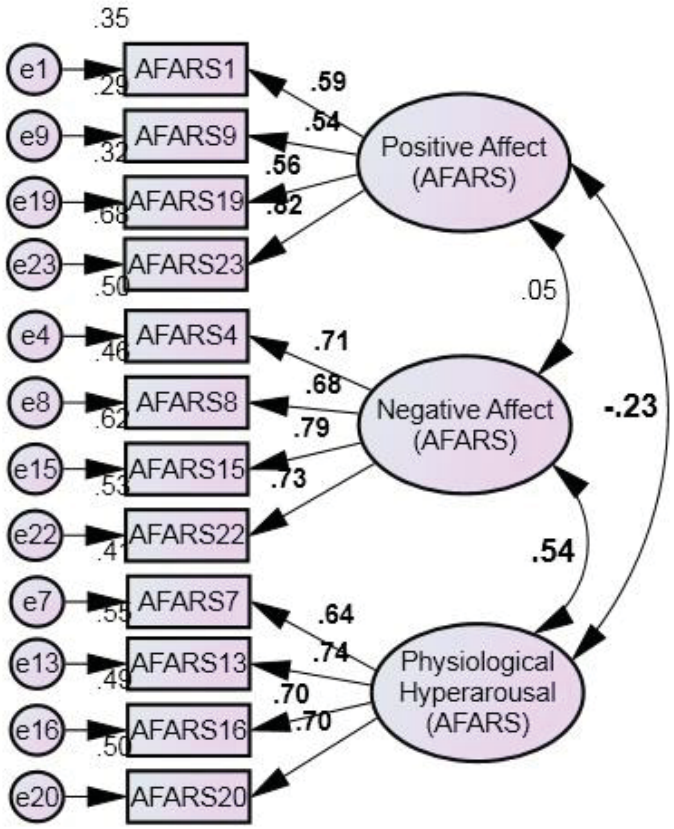


Figure 6.18. Trimmed measurement model of the tripartite factors with AFARS.

As presented in Table 6.9, this model demonstrated good levels of fit across the fit indices. Evaluation of this model using the reduced sample of 188 participants also yielded good to excellent levels of fit across the indices. Comparisons between the trimmed measurement models of the PANAS-C and AFARS appear to indicate AFARS to be of superior fit. Hence, the trimmed measurement model of the AFARS, as presented in figure 6.18, was used to represent the tripartite factors in the structural models. The content of the items used in this measurement model are provided in Appendix G. The resulting PA and PH scales appear to adequately

retain both the breadth of content and the face validity of the original measure. The items used to represent *negative affect* appear to have a narrow focus on feeling upset, for example, "I get upset" and "I get upset by little things". However, this is largely representative of the original scale, as the omitted items also featured ideas of getting upset (e.g., "Other people upset me").

Table 6.9
Goodness-of-Fit Indices for Trimmed Measurement Models of the Tripartite Factors

				Fit Indices						
χ^2	*df*	χ^2/df	*p*	GFI	RMSEA	CFI	TLI	PGFI	PNFI	AIC
Trimmed PANAS-C (*N* = 617)										
141.739	51	2.779	.000	.962	.054	.965	.955	.629	.732	195.739
Trimmed PANAS-C (*N* = 188)										
81.602	51	1.600	.004	.934	.057	.967	.958	.611	.710	135.602
Trimmed AFARS (*N* = 617)										
117.594	51	2.306	.000	.969	.046	.969	.960	.633	.732	171.594
Trimmed AFARS (*N* = 188)										
64.665	51	1.268	.095	.944	.038	.979	.972	.617	.702	118.665

Chapter Seven Results – Structural Modelling

Based on the trimmed measurement models established in the previous chapter, the hypothesised structural relationships between the latent variables were tested. This chapter first examines the structural models involved in the cross-sectional phase of the study, which tests the influence of the tripartite factors on test anxiety while controlling for their influence on depression and anxiety. The resulting structural model from the cross-sectional phase is also tested using the reduced sample for the prospective analyses. The chapter then uses this model as a framework for the prospective analysis of how the test performance is influenced by the tripartite factors of emotions and aspects of test anxiety.

The statistical significance of the regression paths between latent variables was tested based on the benchmark of $p < .05$ and a critical ratio (C.R.) of > +/-1.98 (Byrne, 2010). As discussed in section 5.1.2, it was unclear whether the assumption of independence of cases was met by the prospective sample, and thus the stricter benchmark for statistical significance of $p < .01$ was adopted for the prospective structural models as a precautionary measure. Evaluation of the goodness-of-fit of the structural models employed the same fit indices as for the measurement models. As outlined in Table 4.1, the prospective analysis involved a smaller sample, and thus different benchmarks of goodness-of-fit were adopted.

7.1 Cross-sectional Model of Test Anxiety

A structural model is specified to examine the role of the tripartite factors as precursors to the different aspects of test anxiety, as well as to depression and generalised anxiety. The full structural model being tested is presented in figure 7.1, and represents the integration of the trimmed measurement model of the tripartite factors (figure 6.18), the combined measurement model of depression and generalised anxiety (figure 6.15), and the trimmed multidimensional measurement model of test anxiety (figure 6.10). In the context of the inconclusive literature regarding how gender influences the effects of emotional experiences, gender is placed as a control variable for the dimensions of test anxiety.

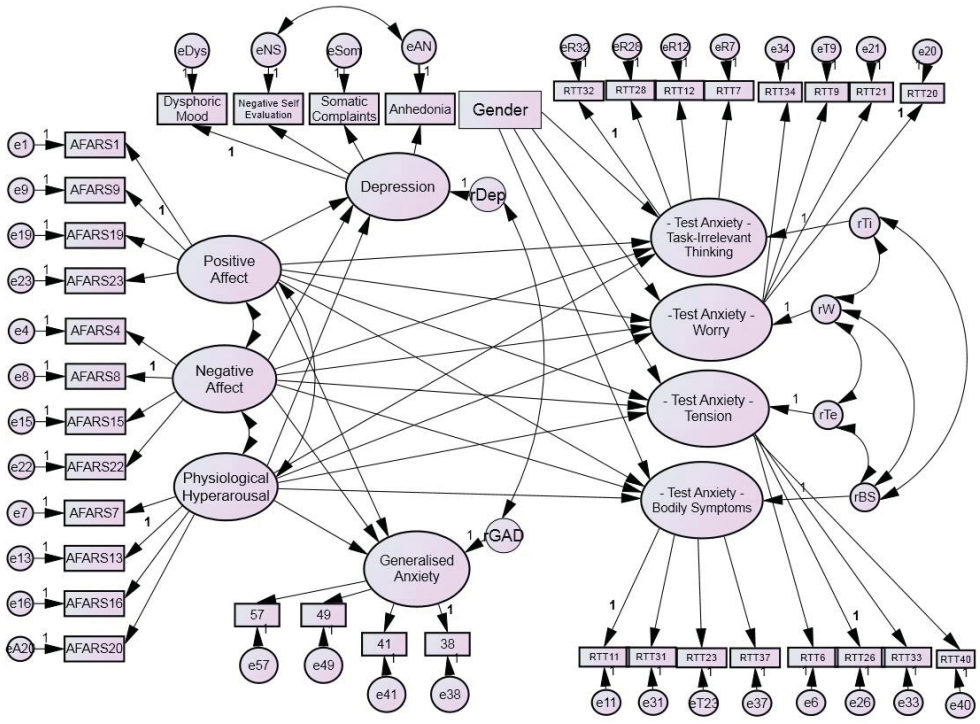


Figure 7.1. Initial structural model for test anxiety.

The resulting model serves to test two of the study's hypotheses. First, it was hypothesised that the influence of the tripartite factors on depression and anxiety is consistent with the original conceptualisation of the tripartite model (i.e., *depression* being influenced by *positive affect* and *negative affect,* and *anxiety* being influenced by *negative affect* and *physiological hyperarousal*). In light of recent findings that challenged the specificity of PA and PH as the distinctive factors for depression and anxiety, respectively, all of the tripartite factors were specified to regress onto both *depression* and *generalised anxiety* in order to explore the strength of these relationships. Second, it was hypothesised that the tripartite factors serve as the emotional precursors to test anxiety. By incorporating aspects of test anxiety as separate endogenous variables along with *depression* and *generalised anxiety,* the model was able to examine the tripartite factors' unique influence on test anxiety, controlling for any features of depression or generalised anxiety that may overlap with test anxiety. The resulting model presented in figure 7.1 is overidentified with 598 degrees of freedom based on 703 data points and 105 freely estimated parameters.

The initial model was tested using the full sample of 617 participants. Figure 7.2 presents the resulting standardised parameter estimates, with bolded values representing those that have met statistical significance. Italicised values above the endogenous latent factors are the squared multiple correlations, which represent the percentage of the factors' variance explained by the exogenous variables. For visual clarity, the manifest variables and their item loadings were omitted from the figure. *Positive affect* (*PA*) significantly influenced depression, but did not have significant effects on generalised anxiety (C.R. = 1.808, p = .071) or any of the four aspects of test anxiety (*worry*: C.R. = -.168, p = .867; *task-irrelevant thinking:* C.R. = -1.640, p = .101; *tension:* C.R. = 1.891, p = .065; *bodily symptoms*: C.R. = .306, p = .759). This is consistent with the proposed role of *positive affect* as the distinctive factor for depression.

The influence of *physiological hyperarousal* (PH) on depression (C.R. = -1.631, p = .103) and generalised anxiety (C.R. = -.492, p = .688) were not statistically significant. Of the four aspects of test anxiety, only *bodily symptoms* was found to be influenced by *PH* (*worry*: C.R. = .425, p = .671; *task-irrelevant thinking:* C.R. = .790, p = .429; *tension:* C.R. = 1.078, p = .281). While this is consistent with the expectation that *PH* does not influence depression, the extent of its influence on anxiety is less than expected, particularly with the non-significant regression to *generalised anxiety.*

The influence of *negative affect* (*NA*) on both *depression* and *generalised anxiety* reached statistical significance, which is consistent with expectations that *NA* is the common factor both these types of emotional disturbance. Its correlation with *PA* did not reach statistical significance (C.R. = .849, p = .396), supporting the view that *PA* and *NA* are orthogonal dimensions. The influence of *NA* was significant for all of the aspects of test anxiety apart from *task-irrelevant thinking* (C.R. = 1.799, p = .072). In fact, the parameter estimates show that *task-irrelevant thinking* was not influenced by any of the tripartite factors. While at the measurement model level *task-irrelevant thinking* was significantly correlated with *bodily symptoms,* the correlation between the residuals of these two factors did not reach statistical significance in the present model (C.R. = .344, p = .731). Apart from its correlation

with the *worry* aspect of anxiety, *task-irrelevant thinking* was shown to be a relatively independent variable from the rest of the model.

The effect of gender on the *tension* and *bodily symptoms* aspects of test anxiety reached statistical significance, suggesting gender differences in the level of physiological symptoms in response to test situations. Gender was not identified to have significant confounding effects on *worry* (C.R. = 1.203, *p* = .229) or *task-irrelevant thinking* (C.R. = -1.789, *p* = .074).

The correlation between the residuals of *depression* and *generalised anxiety* did not reach statistical significance (C.R. = .156, *p* = .876), despite these factors being moderately correlated at the measurement level. This supports the view of the tripartite model of emotions that the relationship between depression and generalised anxiety can be largely accounted for by the influence of the tripartite factors.

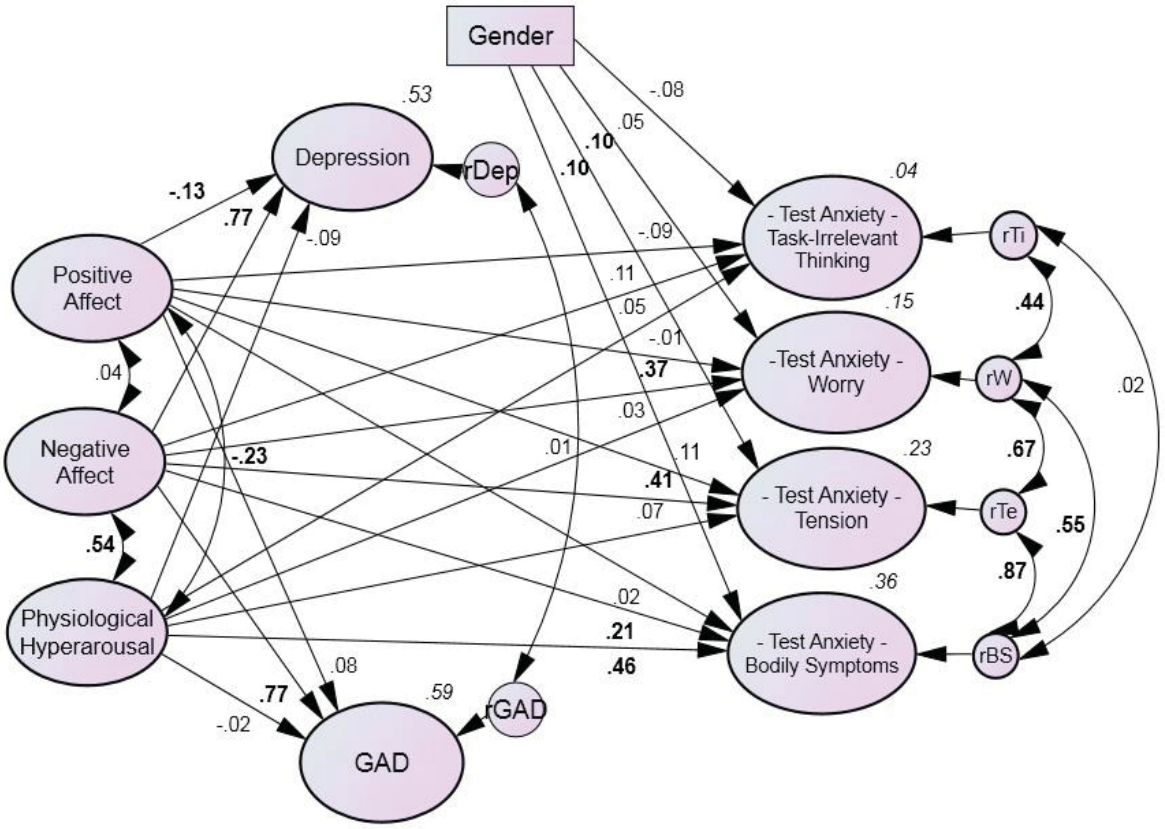


Figure 7.2. Schematic presentation of initial structural model of test anxiety.

Table 7.1 presents with fit indices for Model 1, and reveals that the model did not demonstrate adequate fit across the fit indices. Based on the parameter estimates, non-significant parameters were removed to establish a more parsimonious model (Model 2). Specifically, *NA* was specified as the higher-order construct that influences both *anxiety* and *depression*, as well as the *worry, tension,* and *bodily symptoms* aspects of test anxiety. *PH* was specified to correlate with *NA* and to influence *bodily symptoms.* As *PA* did not demonstrate any meaningful influence in explaining test anxiety, this construct was removed from the model. Similarly, *task-irrelevant thinking* was deemed to be relatively independent from both the tripartite factors as well as other aspects of test anxiety, and was thus removed the model. Gender remains as a control variable for the test anxiety endogenous variables.

Table 7.1
Goodness-of-Fit for the Structural Models of Test Anxiety

				Fit Indices						
χ^2	*df*	χ^2/df	*p*	GFI	RMSEA	CFI	TLI	PGFI	PNFI	AIC
Cross-sectional Model 1 (*N* = 617)										
1520.443	598	2.543	.000	.876	.050	.889	.876	.745	.746	1730.443
Cross-sectional Model 2 (*N* = 617)										
1023.238	364	2.811	.000	.894	.054	.897	.886	.748	.762	1165.238
Cross-sectional Model 3 (*N* = 617)										
447.863	179	2.502	.000	.934	.049	.936	.925	.724	.766	551.863
Cross-sectional Model 3 (*N* = 188)										
268.956	179	1.503	.000	.885	.052	.935	.924	.686	.709	372.956

The resulting structural model (Model 2) and its parameter estimates are shown in Figure 7.3. All regression paths between the latent factors reached statistical significance and were in the expected direction. The regression path of gender on *tension* was statistically significant, suggesting the presence of gender differences in the experience of the *tension* aspect of test anxiety.

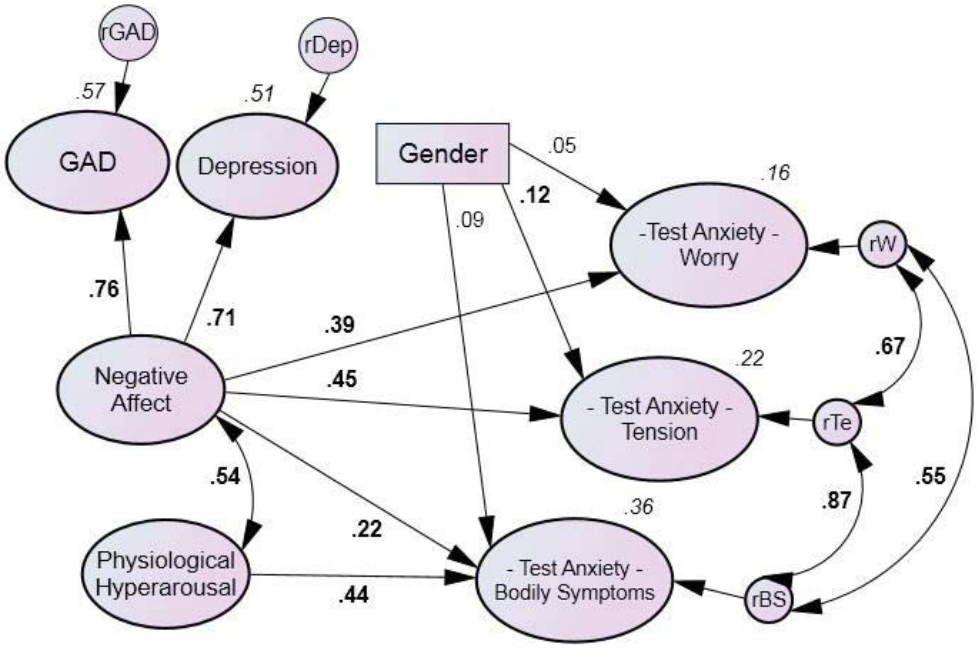


Figure 7.3. Reduced structural model of test anxiety (model 2).

As shown in Table 7.1, Model 2 did not demonstrate adequate fit consistently across the fit indices. Review of the modification indices revealed that model fit may be improved if *NA* and the control variable of gender were specified to be correlated. Substantively, this appears to indicate that there are gender differences in the level of negative affect experienced. This additional correlation path was not specified because it would negate the role of *gender* as a control variable. The list of modification indices also revealed various ways in which *tension* and *generalised anxiety* were closely related. For example, paths with a high modification index included a direct regression path from *tension* to *generalised anxiety,* a correlated path between the residuals of these factors, and the cross-loading of one of the *generalised anxiety* items on to *tension.* The presence of highly related latent variables in a structural model (for example in the case of multicollinearity) may be addressed through removing one of the latent variables and making its manifest indicators become additional indicators for the remaining latent variable – essentially combining the two latent factors into one (Byrne, 2010). This practice was not performed because combining the two factors would inappropriately broaden the construct of test anxiety into non-specific anxiety or distress.

The regression paths for Model 2 suggest that the identified effects of *NA* on test anxiety are beyond those explained by gender differences. Further, having established that *NA* has an influence on test anxiety that is separate from its influence on *depression* and *generalised anxiety*, the latter two constructs can arguably be removed in order to establish a more parsimonious model of test anxiety.

Figure 7.4 presents the final structural model of test anxiety, tested with the full cross-sectional sample. This model is overidentified with 179 degrees of freedom based on 231 data points and 52 freely estimated parameters. As shown in Table 7.1, this model demonstrated adequate levels of fit across the fit indices. This model suggests that a higher level of negative affect is a precursor to higher levels of cognitive worries, apprehensive tension, and physiological symptoms in reaction to testing situations. While the model suggests that there are gender differences in both the *tension* and *bodily symptoms* aspects of test anxiety, the effect of negative affect on test anxiety are beyond those explained by gender effects.

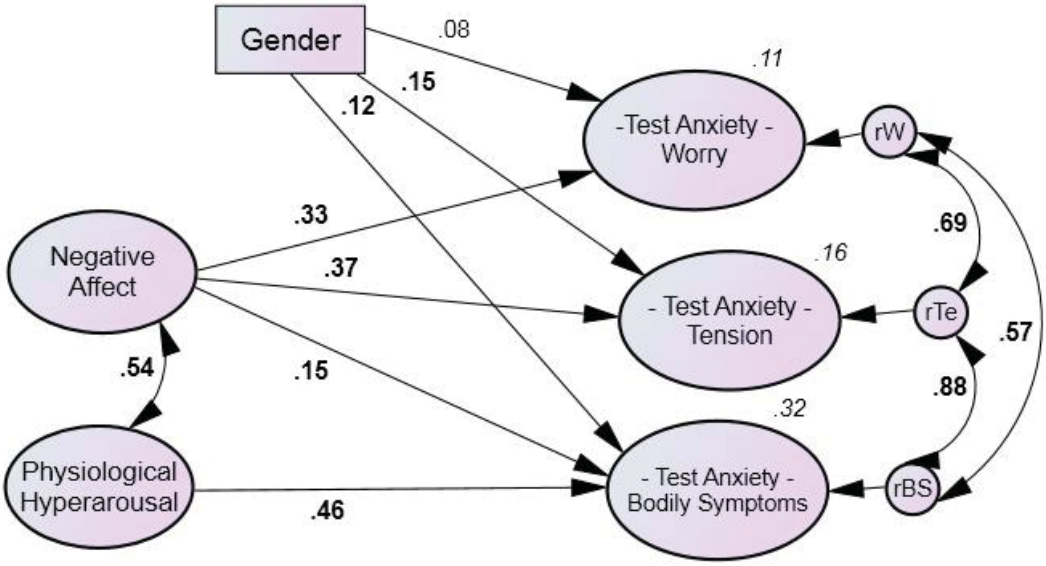


Figure 7.4. Final model (Model 3) of test anxiety for the cross-sectional sample of N = 617.

Model 3 was also tested using the prospective sample of 188 participants in order to verify whether the model remains robust with the reduced sample. As shown in figure 7.5, comparisons of the squared multiple correlations suggest that a slightly

greater amount of the endogenous factors' variance was explained by the exogenous factors for the prospective sample than the cross-sectional sample. However, the influence of *negative affect* on *bodily symptoms* was no longer statistically significant. It is possible that the unique contribution of *negative affect* was masked by the highly correlated residuals between *bodily symptoms* and *tension.* This may also explain the inadequate level of fit as provided by the GFI, suggesting features of model misspecification. Alternative specifications of this model for the prospective sample are explored in conjunction with the added variable of participants' grade point average in order to establish a structural model for academic performance.

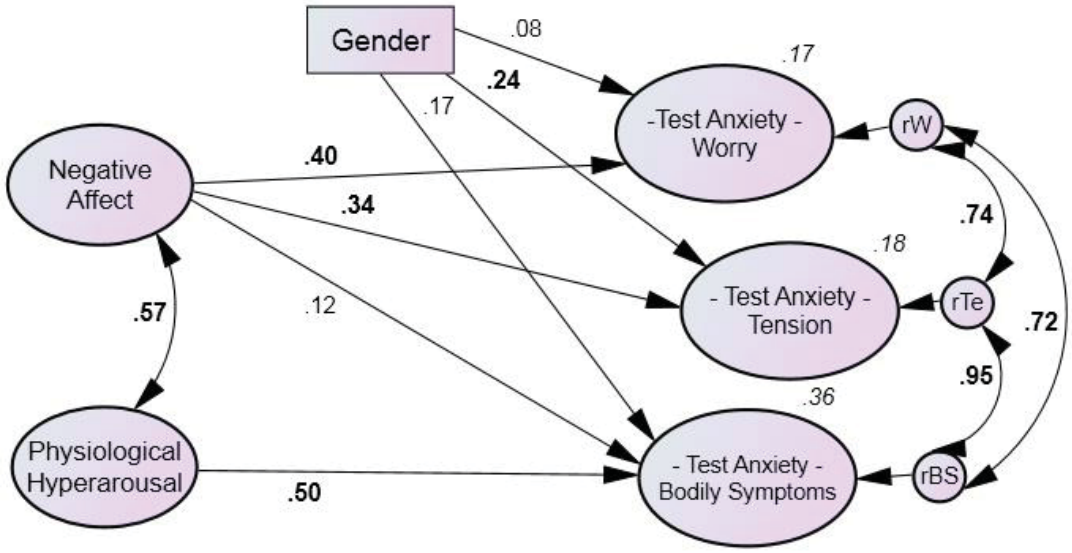


Figure 7.5. Final model (model 3) of test anxiety for the prospective sample of N = 188.

7.2 Prospective Model of Academic Performance

The prospective phase of this study examined the role of negative affect and test anxiety in predicting academic performance. It was hypothesized that the cognitive components of test anxiety would significantly predict academic performance while the emotionality components (i.e., *tension* and *bodily symptoms*) would not. The prospective phase of this study also aimed to reconcile existing research that highlights the indirect effects of emotional distress on academic performance, such as through cognitive appraisals and motivation. As such, the direct and mediating effects of negative affect on academic performance were examined. Using the

previously established structural model for test anxiety (figure 7.5), this section examined the pathways through which general emotional distress may be a meaningful precursor to test anxiety and test performance.

Figure 7.6 depicts the structural model of test anxiety with the additional manifest variable of *grade point average* (GPA). From an initial exploratory approach, both *negative affect* and *physiological hyperarousal,* as well as all three factors of test anxiety, were specified to have direct influences on GPA. The effect of gender on grade point average was also controlled.

As shown in figure 7.6, GPA was not significantly influenced by *worry* (C.R. = -2.164, p = .030), *tension* (C.R. = .691, p = .490), *bodily symptoms* (C.R. = -.449, p = .653), *negative affect* (C.R. = -.030, p = .976), and *physiological hyperarousal* (C.R. = -.354, p = .723). In other words, none of the latent factors appeared to have direct influences on test performance. The standardised regression coefficient for *tension* was greater than 1, which may be indicative of multicollinearity. Together with the r=.95 correlation between the residuals of *tension* and *bodily symptoms,* as well as the similarity in content between these two factors, it was reasonable to assume the presence of multicollinearity.

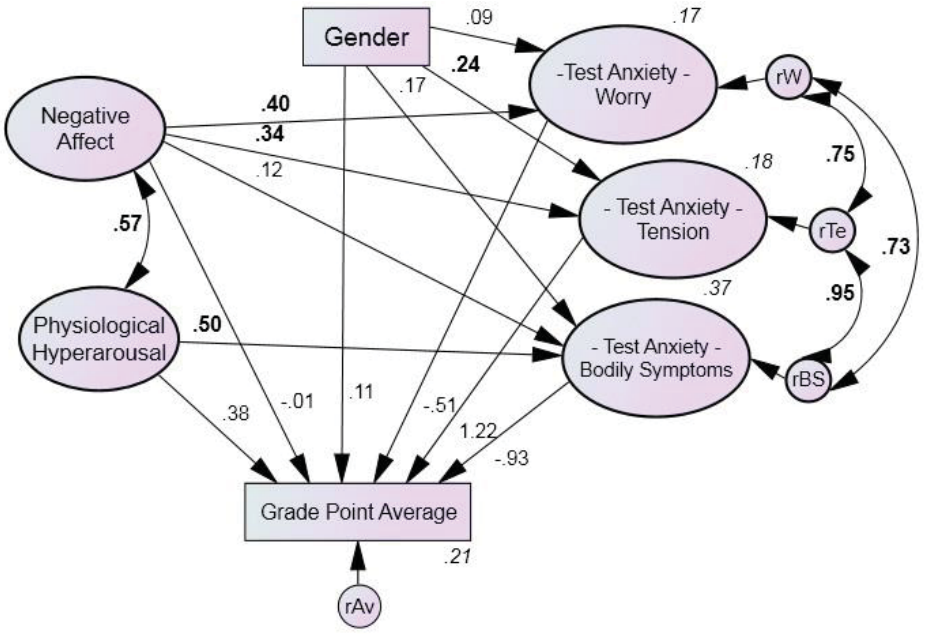


Figure 7.6. Exploratory structural model (model 1) for test performance.

As the *tension* and *bodily symptoms* scales in the RTT were derived from the broader dimension of *emotionality* (Sarason, 1984), it was deemed appropriate to merge these two factors together into a single factor in order to resolve the issue of multicollinearity. Figure 7.7 presents the resulting model where the manifest variables of *bodily symptoms* and *tension* were specified to represent one latent factor of *emotionality* (model 2). It is worth noting, however, that this was carried out at the cost of having a measurement model of test anxiety that was not previously tested. Indeed, as shown in Table 7.2, this model produced poorer levels of fit across all fit indices compared to the original model, suggesting that it was more appropriate for *tension* and *bodily symptoms* to remain as separate constructs.

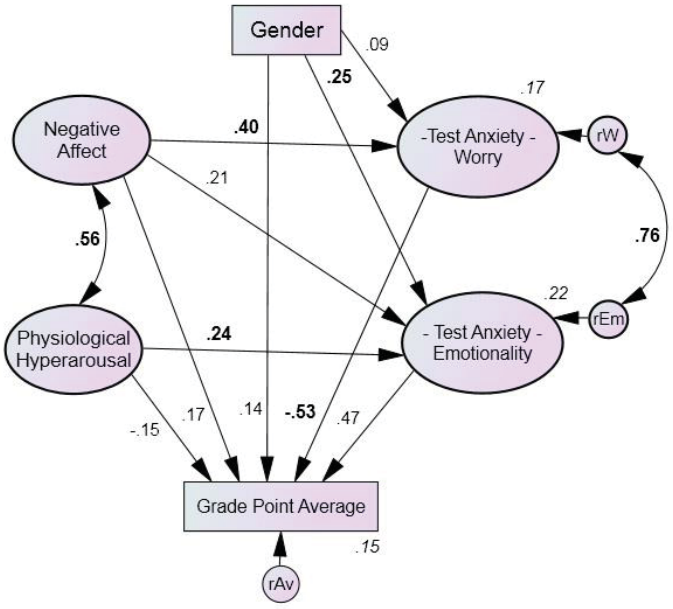


Figure 7.7. Alternative structural model (model 2) of test performance.

Table 7.2
Goodness-of-Fit for the Structural Models of Test Performance

				Fit Indices						
χ^2	*df*	χ^2/df	*p*	GFI	RMSEA	CFI	TLI	PGFI	PNFI	AIC
Prospective Model 1 (*N* = 188)										
292.961	194	1.510	.000	.880	.052	.930	.917	.675	.690	410.961
Prospective Model 2 (*N* = 188)										
314.051	199	1.578	.000	.872	.056	.919	.906	.686	.697	422.051
Prospective Model 3 (*N* = 188)										
191.635	125	1.533	.000	.900	.053	.944	.932	.658	.701	283.635
Prospective Model 4 (*N* = 188)										
120.445	73	1.650	.000	.915	.059	.950	.938	.636	.709	184.445

As a further attempt to resolve the apparent redundancy of contents identified in model 1, an alternative approach was used where the latent factor of *bodily symptoms* was removed from the model, along with its manifest indicators. While this served to address the presence of multicollinearity, the resulting model (model 3; figure 7.8) had the disadvantage of being less inclusive of the construct of test anxiety, having only two of the four aspects of test anxiety represented. This model produced adequate levels of fit across the fit indices apart from the GFI (.900), which is short of the benchmark of >.920. The regression coefficients for this model reveal that *negative affect* positively influences both the *worry* and *tension* aspects of test anxiety, accounting for approximately 17% of their variance. GPA was not predicted by *NA* (C.R. = 1.411, *p* = .158) or *PH* (C.R. = -1.058, *p* = .290), and the effect of gender on GPA did not reach statistical significance (C.R. = 1.653, *p* = .098). There was a statistically significant gender effect for *tension*, but not for *worry* (C.R. = 1.110, *p* = .267). *Worry* had a negative influence on GPA, while *tension* demonstrated a positive influence on GPA.

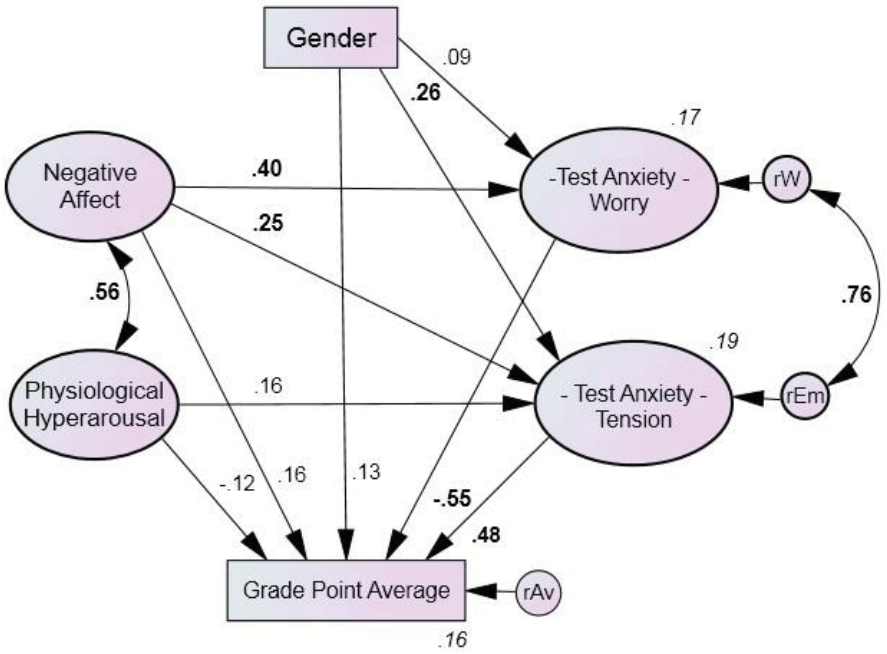


Figure 7.8. Alternative structural model (model 3) of test performance.

A final parsimonious model of test performance was specified by removing the non-significant regression paths in model 3. The latent variable of *PH* was also removed because it did not demonstrate meaningful direct or indirect influences on GPA or test anxiety, nor was it deemed theoretically relevant to include *PH* in the model as a control variable.

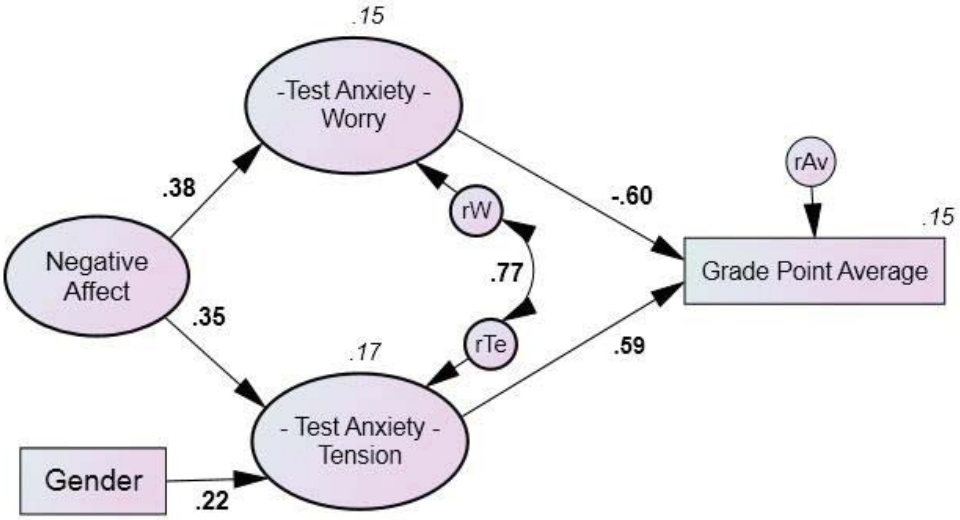


Figure 7.9. Final structural model (model 4) for test performance.

Model 4 appears to indicate *negative affect* to have indirect effects on test performance that are mediated through the two aspects of test anxiety. Reconciling recommendations by Baron and Kenny (1986) and MacKinnon, Lockwood, Hoffman, West, and Sheets (2002), such inferences on mediation were tested through three steps of analyses.

First, the effects of the exogenous variable (NA) on each of the mediating variables were separately tested, while controlling for the effects of gender. For example, the unstandardised regression coefficient for NA → Worry, as presented in Table 7.3, was calculated from a model with *negative affect* and *gender* as the exogenous variables, each with a regression path to *worry.* A similar model was specified for *tension* to establish the regression coefficient for NA → Tension. The regression coefficients indicate that higher levels of *worry* and *tension* were predicted by higher levels of *negative affect.*

The second step was to establish the influence of the mediating variables on the outcome variable of GPA, while controlling of the influence of gender and the predictor variable of *negative affect*. For example, the coefficient for Worry → GPA, as presented in Table 7.3, was calculated from a model with *worry, negative affect,* and gender having direct regression paths to GPA. Higher GPA was predicted by lower levels of *worry*, thus indicating a possible mediation path for *negative affect* on GPA. The influence of *tension* on GPA, however, was not statistically significant, as it was confounded by the effect of gender (β = .347, $p <$ .01).

Table 7.3
Tests of Conditions for Mediation

Regression paths	Unstandardised regression coefficient (*p*)
NA → Worry	.405 (<.01)
NA → Tension	.441 (<.01)
Worry → GPA	-.211 (.03)
Tension → GPA	.052 (.49)

NA: Negative affect. GPA: Grade point average.

The third step examined the indirect effect of *negative affect* on GPA via *worry,* which was carried out through bootstrapping using the AMOS software. Regression paths were specified from *negative affect* to *worry*, and from *worry* to GPA. The effects of *tension* and gender were controlled through direct regression paths from these variables to GPA. The bootstrap analysis generated 95% confidence intervals for the indirect path, and can indicate a significant mediation effect if the upper and lower bounds do not cross zero, along with a bias-corrected $p < .05$ (Fritz & MacKinnon, 2007; MacKinnon et al., 2002). The resulting upper and lower bounds for the unstandardised indirect effect were -.029 and -.292, respectively, at $p =$ .014. Fritz and MacKinnon (2007) provides the required sample size for .80 power to detect mediation effects, which is estimated based on the type of analysis performed and the strength of the direct regression paths along the mediation chain. The recommended sample size for bias-corrected bootstrap medium sized regression coefficients was 116 (Fritz & MacKinnon, 2007), suggesting that the

current sample of 188 was sufficient to detect the mediation effects. These findings strongly suggest that negative affect has an indirect negative influence on GPA that is completed mediated by worry. However, this relationship was short of reaching the probability benchmark of $p < .01$.

The indirect effect of *negative affect* via *tension* was also tested. The resulting upper and lower bounds for the unstandardised indirect effect were .263 and .025, respectively, at $p = .014$. This indicates that the mediation effect was statistically significant, and was also identified in conjunction with a statistically significant confounding influence of gender ($\beta = .330$, $p < .01$). However, based on Fritz and MacKinnon's (2007) estimates, evaluation of this mediation effect would require a sample size of 391, which is not met by the current study's sample. Together with the non-significant regression path from *tension* to GPA, as identified in Table 7.3, as well as the significant confound of gender effects, there is insufficient evidence to suggest a mediation effect for *tension.* Hence, while the final structural model for test performance (figure 7.9) illustrates a potentially positive indirect effect on negative affect on GPA, this relationship could not be reliability substantiated statistically due to insufficient sample size and statistical power.

In summary, the prospective examination into the effects of test anxiety on test performance has replicated the existing finding that the cognitive aspects of test anxiety significantly and negatively affect test performance. Contrary to expectations, however, the effect of *tension* was positive and reached statistical significance. This final model accounted for 15% of the variance in exam grades, which exceeded the level identified in previous studies (Cassady & Johnson, 2002; Chapell et al., 2005; Hembree, 1988). However, the role of negative affect was unclear. While it reliability predicted the *worry* and *tension* aspects of test anxiety, its direct and indirect effects on test performance could not be definitively established or ruled out due to several factors, such as the selection of the significance level for *p,* and the insufficient sample size.

Chapter Eight Discussion

The present investigation involved two main phases that examined the relationships between the tripartite factors of emotional experiences, test anxiety, and test performance. This chapter discusses the findings across the cross-sectional and prospective phases of the analysis, and reviews how the study's hypotheses were largely supported. The relevance of the findings to the clinical and educational fields are highlighted, but also discussed in the context of the various limitations of the study.

8.1 Cross-Sectional Study – Tripartite Model of Emotions

The cross-sectional structural model for test anxiety attempted to determine whether the tripartite factors of general emotional distress have direct effects on test anxiety that are beyond their known influence on depression and anxiety. The sequence of alterations to the structural model for test anxiety (figure 7.2) suggests that negative affect and physiological hyperarousal directly predict certain aspects of test anxiety. Specifically, *negative affect* directly influenced the *worry*, *tension*, and *bodily symptoms* aspects of test anxiety, while PH influenced only the *bodily symptoms* aspect. Because separate endogenous variables of depression and generalised anxiety were specified, the above relationships arguably represented the factors true influence the construct of test anxiety, rather than depressive or anxiety features. Also, these relationships were significant beyond the identified confounding gender effects for *tension* and *bodily symptoms.* All these findings indicate that negative affect and physiological hyperarousal reliably predict certain aspects of test anxiety, providing support for the first hypothesis.

However, the way that construct of test anxiety was representing in the final structural model differed from its original conceptualisation, namely surrounding the role of *task-irrelevant thinking.* The task-irrelevant thinking aspect of test anxiety was not influenced by any of the tripartite factors of emotional distress. As demonstrated in the measurement model of the Reactions to Test scores, task-irrelevant thinking had a comparatively low factor loading onto the broader

concept of test anxiety, and was also correlated with the other three factors to a weaker degree compared to the other factors. These findings suggest that task-irrelevant thinking may not be highly indicative of the general concept of test anxiety, but rather a stand-alone indicator for the tendency to be distractible. This is consistent with Nicholson's (1958) review of drive theory which views task-irrelevant responses as the mechanism through which anxiety influences performance, rather than a component of anxiety itself. Similarly, Keogh, Bond, French, Richards, and Davis (2004) examined distractibility as having a separate and mediating role between test-related worries and test performance. As task-irrelevant thinking was captured by items that largely represent distractibility (e.g., "I daydream during tests" and "During tests I think about recent past events"), parallels may be drawn to argue that the latent variable of *task-irrelevant thinking* was separate from the broader construct of test anxiety.

As such, it was deemed appropriate to remove that factor from the final structural model of test anxiety. The relative disadvantage was that the influence of *task-irrelevant thinking* on test performance could not be examined at the prospective structural modelling level. As indicated in Table 5.7, task-irrelevant thinking had a statistically significant birvariate correlation with grade point average, and it is possible that this relationship can be identified at the multivariate level. However, given that task-irrelevant thinking was argued to be separate from the construct of test anxiety, analysis of its influence on test performance will not provide meaningful information to the present research questions. With sufficient theoretical justification (e.g., Keogh et al., 2004), future analysis can examine the potential confounding or mediating effects of task-irrelevant thinking on test performance.

The latent variable of positive affect in the initial structural model (model 1) was not shown to have any meaningful influence on test anxiety. However, this initial model did not produce adequate fit to the sample data, and the latent variable of positive affect was removed from the model before an adequately fitting model could be established. Hence, the role of positive affect could not definitively be ruled out. Nevertheless, the finding that only negative affect and physiological

hyperarousal, but not positive affect, influenced aspects of test anxiety provided further support for the utility of the tripartite dimensions in differentiating anxiety from depressive mood disturbances.

As previously mentioned, the cross-sectional phase examined how the tripartite factors of general emotional experiences may predict levels of test anxiety, while controlling for their influence on depression and generalised anxiety. Thus at the same time, the cross-sectional phase also provided findings into the applicability of the tripartite model of emotions (TME; Clark & Watson, 1991) by examining how the tripartite factors influence depression and generalised anxiety. It was predicted that the experience of depressive and anxiety symptoms would be sufficiently differentiated for the current sample of older adolescents, insofar as they are influenced by different combinations of the tripartite factors. In other words, similar to findings for adult populations, positive affect (PA; or lack thereof) and negative affect (NA) influences greater levels of depressive symptoms, whereas NA and physiological hyperarousal (PH) influences anxiety symptoms. Also of interest was how the three tripartite dimensions were inter-related. The following paragraphs review the observed relationships between three tripartite dimensions, and subsequently how these factors influenced depression and generalised anxiety.

The final measurement model of the tripartite factors (figure 6.17), as well as the initial structural model of test anxiety (figure 7.2), revealed an orthogonal relationship between PA and NA. This is consistent with Clark and Watson's (1991) original assertions that they are independent affective dimensions rather than the opposite ends of a unidimensional spectrum. However, extant research has revealed inconsistent findings regarding whether the tripartite dimensions are interrelated or independent constructs (e.g., Joiner et al., 1996; Laurent et al., 2011; Lonigan et al., 2003). Such inconsistencies are in part due to the lack of specific measures for the tripartite dimensions, and the observed correlations reflected the measures' lack of discriminant validity (Anderson & Hope, 2008). While the development of the PH-PANAS-C and AFARS enabled these relationships to be studied with greater construct validity, it appears that measurement factors continue to influence the observed relationship between the tripartite dimensions.

For example, pervasive patterns are seen when comparing the present study's intercorrelations among the PH-PANAS-C and AFARS subscales with those of three other studies that have similarly administered both these tools with child or adolescent samples (Chorpita & Daleidin, 2002; Jacques & Mash, 2004; Stevanovic, Laurent, & Lakic, 2013). The direction and magnitude of the between- and within-measure intercorrelations for the present study were largely comparable to the other studies (see Appendix H). Review of the between-measure correlations reveal generally low to moderate correlations, from r = .34 to .53 between the two PA scales, and from r = .36 to .63 between the two NA scales. This pattern suggests that the two measures may be capturing different aspects of the tripartite dimensions (e.g., Chorpita & Daleidin, 2002). Further, there is an apparent pattern where the PA-NA correlations within the PH-PANAS-C measure were statistically significant across all four studies, whereas the PA-NA correlations for the AFARS were consistently of a lower magnitude and did not always reach statistical significance. Together, the patterns identified in these studies illustrate persistent differences in the measures' psychometric qualities, which likely contribute to the various inconsistent findings in the literature such as the orthogonality of the PA and NA dimensions. The analyses in the present study revealed that the AFARS measure produced superior measurement model of the tripartite dimensions than the PH-PANAS-C, and hence would support the factor structure of the AFARS which shows PA and NA to be statistically orthogonal.

In terms of how the tripartite dimensions influence to depression and anxiety, support was found for the role of NA as a common underlying factor for depression and generalised anxiety. Out of the three tripartite factors, NA was most strongly related with depression and anxiety scores at a bivariate correlation level and at the multivariate structural level. Consistent with other similar investigations, these results imply that the experience of broad negative emotional states is a predisposing factor for more specific symptoms of depression and anxiety, and may serve to account for the high rate of co-morbidity between these disorders (Chorpita Daleiden et al., 2000; Clark et al., 1995; Watson, Gamez, & Simms, 2005).

At a correlational level, the statistical significance of PA's association with depression and anxiety was sensitive to the measure used, further highlighting the difference in the utility between the PH-PANAS-C and the AFARS (Chorpita & Daleiden, 2002; Jacques & Mash, 2004). Across both measures, PA was significantly associated with depression, but also with social phobia and school phobia, which is inconsistent with the TME. Such relationships between low PA and social anxiety have been frequently observed in previous studies with clinical child (Anderson et al., 2010; Chorpita, Plummer et al., 2000) and adult samples (T. Brown et al., 1998; Watson, Clark, & Tellegen, 1988). Alternative conceptualisations of PA as a personality trait were also found to be negatively associated with social anxiety symptoms for adult samples (Naragon-Gainey, Watson, & Markon, 2009; Sellbom, Ben-Porath, & Bagby, 2008; Watson et al., 2005). Hence, the finding that PA may account for the shared variance between depression and social anxiety appears replicable across populations and theoretical models of emotions, which reflects the role of distress in interpersonal situations in underlying the impairment in social functioning across these disorders (Anderson & Hope, 2008).

Due to inadequacies in a multi-factor measurement model of anxiety, the present study elected to represent the concept of anxiety using only one type of anxiety, namely *generalised anxiety*. This resulted in a significant narrowing of content that is represented in the structural model. Nevertheless, at this structural level, PA did not significantly influence generalised anxiety, but its influence on depression reached statistical significance. However, the magnitude of such influences suggests that its substantive role as a precursor to depressive symptoms may be limited for the current sample. Together with PA's lack of influence on any of the aspects of test anxiety, the findings suggest that the experience of general positive emotions does not serve a meaningful protective role against the experience of test-related emotional distress.

In terms of PH, both PH-PANAS-C and AFARS PH scores were significantly correlated with all depression and anxiety subscales scores, suggesting that it may play a similar role to NA as a general distress factor. This is inconsistent with the tripartite model which specifies PH as the distinctive factor for anxiety, and with other

research that views PH as the specific indicator for anxiety, in particular panic disorder (e.g., Kring et al., 2007; Mineka et al., 1998). The relationship between PH and depression was also identified in previous studies with clinical and non-clinical samples of children and adolescents (Chorpita & Daleiden, 2002; de Bolle et al., 2010; Jacques & Mash, 2004), which possibly reflects how depression may be experienced predominantly through somatic symptoms for this age group (APA, 2000; Carr, 2006). However, in the present study, the relationship was not found to be significant at the structural modelling level, where the unique effect of PH was examined independently from the influence of other variables. Combined with the finding that PH was associated with NA, which had a significant influence on depression, the results suggest that PH's bivariate correlation with depression was largely due to its relationship with NA (Anderson & Hope, 2008).

8.2 Prospective Model

Subsequent to establishing a structural model for test anxiety, the prospective phase examined the model's utility in predicting academic performance. The initial structural model for test performance was identified to be adequately fitting on all indices apart from one (GFI). Prior to the removal of *bodily symptoms* form the model, this initial model indicated that test performance was not influenced by the various aspects of test anxiety. As previously discussed, this was likely due to the presence of multicollinearity, where the overlap between the aspects were so substantial that each of the factors' unique influences on test performance could not be detected. Indeed, this problem appears to be resolved through removing one of the factors of test anxiety.

Other explanations of the non-significant effects of test anxiety should also be considered. For example, consideration should be given to the nature of test anxiety that was captured by the study's questionnaire. The Reactions to Tests measure was intended to capture test anxiety as a state construct, with items that ask how respondents feel "before tests" or "during tests". As such, inferences about the role of test anxiety are made with the assumption that the data reflected the actual level of emotional distress experienced during testing situations. This is

important because the theoretical explanations on how anxiety impairs performance, such as the cognitive interference theory (Sarason, 1984), are based on anxiety states. In practice, the questionnaire more likely captured the participants' retrospective estimation of how they usually fell towards testing situations, which is more synonymous with a trait measure. Also, the questionnaire was administered approximately three months prior to their end-of-year examinations. Thus it is unlikely that the participants were experiencing the same level of test anxiety as they would during the examination period, nor would they have necessarily accurately remembered how they felt during the previous year's exams. It is likely that true measures of state test anxiety, such as administering brief measures immediately prior to an exam or the use of physiological monitors during the exam, may yield more robust estimates of their influence on test performance.

The final structural model replicated the existing finding that the cognitive aspects of test anxiety significantly and negatively affect test performance (Cassady & Johnson, 2002; Deffenbacher, 1980; Hembree, 1988). Contrary to expectations, however, the effect of *tension* reached statistical significance and was positive in direction, suggesting that higher tension leads to higher exam grades. The tension subscale of the RTT reflects the general sense of apprehension towards exams, with items such as "I feel distressed and uneasy before tests" and "I feel panicky during tests". It is possible that the finding reflected the facilitative role of anxiety where mild levels of apprehension activate greater cognitive functioning and performance (Alpert & Haber, 1960; Hembree, 1988). Indeed, a parallel may be drawn with Lowe et al.'s (2008) findings that performance enhancing/facilitation anxiety (e.g., "I do better on a test if I am slightly worried before I take it") was weakly associated with better performance in arithmetic tasks.

It is noteworthy that the present analysis estimated tension to have a linear relationship with exam performance, whereas the threshold effect of facilitative anxiety may reflect a curvilinear relationship where anxiety only improves performance when it is below a certain level of severity (Fletcher et al., 1997).

Further clarification of this finding may require explicit comparisons between low-anxious and high-anxious groups.

The final structural model for test performance demonstrated how the *worry* and *tension* aspects of test anxiety accounted for 15% of the variance in exam grades, which exceeded the level identified in previous studies (Cassady & Johnson, 2002; Chapell et al., 2005; Hembree, 1988). Pekrun (2006) asserts that test anxiety is one of many emotions that may be experienced before a test, for example due to the student's simultaneous appraisal of the chance of success as well as threat of failure. It is anticipated that incorporating more specific types of emotional experiences, rather than a general construct of negative affect, will help explain a greater proportion of the variance in test performance.

As the direct effect of negative affect on GPA was not identified to be significant, its statistical contribution to explaining the variance in test performance would be negligible. A possible explanation for the non-significant direct influence of NA on test performance is that the construct may be too broad, and hence lack the specificity required to distinguish between activating and deactivating types of emotional distress. For example, anticipatory anxiety and shame may be associated with a greater sense of control or personal responsibility, hence become activating agents for facilitating achievement. On the other hand, negative emotions associated with a low sense of personal control, such as boredom and hopelessness, become deactivating agents with a deleterious effect on achievement (Daniels et al., 2009; Pekrun, 2006).

In the present study, 'negative emotions' was seen as a unitary concept, represented by items that largely captured participants' tendency to 'feel upset'. As such, it likely masked the different mechanisms through it may both positively and negatively influence test performance. Indeed, only through more specific mediating variables do the indirect effects of NA become apparent. Specifically, NA directly predicted the worry aspect of test anxiety which impaired exam grades, and also influenced the tension aspect of test anxiety which served to facilitate better performance. The present findings support the need to identify more specific

cognitive and emotional states in order to understand the influence of negative affect on test performance. More broadly, the present findings support the view that experiencing general emotional distress – at non-clinical levels – does not necessarily impair test performance (Lowe et al., 2008; Pekrun, 2006).

A second explanation relates to the inadequate sample size for the prospective phase of the study and how the resulting lack of statistical power accompanies the increased propensity for Type-II error. Based on the extant research base, the relationships among the present study's variables were expected to be of a small to medium effect size (e.g., Cassady & Johnson, 2002; Hembree, 1988; King et al., 1995; Seipp, 1991), which require larger sample groups in order to achieve sufficient statistical power to reliably detect the effects. Further, the method of structural equation modelling has high and unique demands on sample size, which are proportional to the number of observed and latent variables in the model (Hair et al., 2006). While the study's sample was arguably sufficient to reliability test the reduced final cross-sectional model for test anxiety and the final prospective model for test performance, the sample size was insufficient to detect potential indirect effects of negative affect on test performance.

8.3 Implications for Practice

The TME provides a model for understanding the similarities and differences between anxiety and mood disorders, and has several implications for clinical practice. Researchers have recommended clinicians to examine the distinguishing factors of PA and PH to aid with differential diagnosis, or to inform treatment planning for individuals with co-morbid depressive and anxious symptoms (Tully et al., 2009). However, the present study raises caution for such utility, as the magnitude of PA's influence on depression was small, and PH's influence on generalised anxiety was not statistically significant. It is also important to acknowledge that the level of depressive or anxiety symptoms experienced by the current community-based sample of students would not be deemed clinically significant, or equivalent to those by individuals diagnosed with these disorders. Hence, the present study has reservations in making inferences on how the findings

may inform the differential diagnosis of mood and anxiety disorders. Nevertheless, the findings demonstrate how the influence of negative affect appears pervasive across other types of emotional or test-specific distress. This supports the rationale for using trans-diagnostic therapy models or strategies that focus on managing the underlying emotional distress, without necessarily needing to adopt separate treatment protocols for depression and anxiety problems.

Teachers who are observant of their students may readily recognise those who are experiencing general emotional distress. However, such observations often do not reveal the specific type of negative emotion that is being experienced, and it can be difficult to differentiate, for example, boredom from hopelessness, or shame from guilt. Further, students' internal thought processes, such as negative appraisals of the test outcome or other test-specific worries, may be even more difficult to observe. The present study contributes to the question of whether in-depth exploration into students' specific emotional experiences and thought processes is useful (if not crucial) for informing ways of intervention. The findings suggest that the answer depends on the desired goal of intervention. Specifically, if the goal of intervention is to reduce emotional distress or improve the student's overall emotional wellbeing, the focus on general negative affect as a therapeutic target may be sufficient. That is because negative affect serves as an important precursor to the experience of anxiety and depressive symptoms, as well as the tendency to experience cognitive and emotional aspects of test anxiety. This suggests that trans-diagnostic interventions for building resilience towards general emotional distress will reap therapeutic benefits across other specific types of emotional distress.

On the other hand, in-depth assessment or exploration of specific forms of distress is recommended if the goal of intervention is to reduce the negative impact of emotional distress on test performance. Negative affect, or the general tendency for the individual to feel upset, was not identified to have a direct effect on test performance. However, when mediating conditions such as levels of test anxiety were specified, the indirect effects of negative affect on test performance were revealed. Thus the impact of general negative affect on test performance can only

be understood in the context of more specific types of emotional distress. This is consistent with the Control-Value theory of achievement emotions, which distinguishes between activating and deactivating types of negative emotional states and associated cognitive appraisals (Daniels et al., 2009; Pekrun, 2006). Based on the present findings, it is recommended that teachers or school-based clinicians who work with emotionally distressed students should screen for specific worries towards examinations, including any elevated perceived threat of test difficulty, underestimation of self-ability, and the consequences of failing. These worries were shown to have detrimental effects on exam performance. In the absence of such worries, the students may be given some reassurance that the experience of anticipatory tension is normal, if not potentially beneficial for improving test performance.

It is again noteworthy that the present study investigated the influence of the negative affect on one specific outcome variable of examination GPA. It will be premature to rule out how negative affect may influence more general aspects of academic functioning, such as course-work completion, engagement with learning, or class attendance (Hughes et al., 2008; Keogh et al., 2004; Pekrun, 2006; van Ameringen et al., 2003). Similarly, despite how the present findings indicated *positive affect* to have limited influence on test anxiety, the importance of promoting students' positive affect should not be undervalued. Indeed, research into school-based programmes for promoting positive emotions in schools had yielded favourable outcomes across many indicators of psychological wellbeing (Ruini et al., 2009).

8.4 Strengths, Limitations, and Future Directions

The measurement of the constructs was a merit of this study. The tripartite factors of positive affect (PA), negative affect (NA), and physiological hyperarousal (PH) were measured using the PH-PANAS-C and the AFARS, which were specifically designed to capture these concepts in their pure affective form. This represents an advance from early studies of the tripartite model which used measures of depression and anxiety to capture the tripartite factors (e.g., Cole et al., 1997;

Lonigan et al., 1999), and was considered an important step forward in this area of research (Anderson & Hope, 2008; de Bolle & de Fruyt, 2010). The measures of depression and anxiety demonstrated superior discriminant validity than those used by Jacques and Mash (2004), suggesting that this study was more able to capture these disorders as separate symptom groups. Specifically, the correlation of r = .55 between the RADS-II and SCARED-R total scores in the present study was lower than the correlation of r = .71 between the CDI and STAI-C observed by Jacques and Mash.

It is noteworthy that all data (apart from exam grades) were gathered through self-report measures, and thus there was a reliance on the participants' insight into their own emotional experiences and reactions. However, the use of self-report in the absence of collaborative information appears supported. Previous research had complemented self-report with parental-report measures of emotional distress, but found that there were discrepancies between the sources of information (Daleiden et al., 2000). Given that the tripartite dimensions represent higher-order constructs of internal experiences that are not always observable, it was argued that they are more accurately captured through self-report measures, and that collaborative information from parental-reports provides little additional information (Anderson & Hope, 2008; Daleiden et al., 2000).

A problematic area of measurement in the present study was the calculation of participants' grade-point-average (GPA) as the sole indicator of test performance. Following consultation with a school principal, it became apparent that several features of the NCEA grading system made it difficult to establish a consistent indicator of test performance across all participants (M. Scott, personal communication, July 20, 2010), and readers should keep the following sources of variation in mind when interpreting the results of this study. Firstly, there are elements of strategic planning involved in completing the exams. The overall goal for some students may be to pass only a certain number of exams in order to attain enough credits to progress their qualification. As such, it is at times a useful strategy for students to allocate their time to focus only on a selection of exams, and choose to ignore or '*non-submit*' others. From this perspective, an average

grade may be a problematic indicator for comparing students' performance. For example, Student A who attempted to complete all four exams in Chemistry but failed one of them would have a lower average score (e.g., 3x Achieved (6) + 1x Not Achieved (0) / 4 = 1.5) compared to Student B who only attempted three exams (e.g., 3x Achieved (6) / 3 = 2.00). It is logically problematic to determine student B as having superior performance when in fact the higher GPA reflected the strategic non-submission of an exam.

Beyond exam-taking strategies, there is an inherent variation in the number of exams undertaken by each participant due to differences between academic subjects and year levels. The number of exams taken by the participants ranged from 2 to 25, and the researcher did not collect any information (e.g., from interviewing each participant) that would establish the reason for each variation. Due to these potential confounds, the variations in GPA between participants may not reflect genuine differences between their ability to perform in exam situations.

Another problematic issue in terms of capturing the study's variables was that the initial specifications of the study's measurement models and structural models all indicated inadequate levels of fit, and thus alternative trimmed models were tested. The process of reducing the structural models had led to several variables being removed from the structural modelling analyses. Namely, the analysis was not able to establish whether test performance was predicted by *task-irrelevant thinking* or by *positive affect*. These factors were significantly related to test performance at the bivariate correlation level, and thus there is potential for their influence to be examined at the multivariate level. Based on the present findings, future studies should consider specifying *task-irrelevant thinking* as separate from the concept of test anxiety, and perhaps as more synonymous to *distractibility* (e.g., Keogh et al., 2004).

Another feature of the present study that warrants further exploration is the differential effect of emotional distress across different academic subjects. Justification of such an investigation is apparent in light of the literature on *maths anxiety* (Ma & Xu, 2004), which has been shown to have a negative influence on

test performance beyond that accounted for by general test anxiety (Devine et al., 2012). Further, reductions in test anxiety were found to produce greater improvements in maths test performance compared to that of English or science and technology subjects (Gregor, 2005). Parallel to this, research within the cognitive psychology framework demonstrated that different subject tasks may be influenced by different mechanisms of working memory. For example, it was argued that mathematics required greater *cognitive flexibility*, while comprehension-based subjects such as reading and social studies required greater *monitoring* or evaluation of information (Bull et al., 2008; Latzman et al., 2010). These findings highlight a degree of heterogeneity among different subject demands in terms of its tendency to be influenced by emotional distress.

In the present study, the participants' academic performance across all subjects was subsumed into one GPA, thus possibly masking any substantial relationships between emotional distress and a particular type of subject demand. Given that this study substantiated the indirect influence of *NA* on test performance, it invites future theoretical development and empirical testing into the role of emotional distress across specific subject demands. The use of exam GPA as the only indicator of test performance reflected the focus on the interference effects of emotional disturbance on the test-taking situation, which was in line with the aims of this study. However, this narrow view of academic performance inevitably neglects other deficit and motivational indicators of academic functioning, such as the completion of and performance on homework assignments, efficiency of learning, study habits, school attendance, or attainment of qualifications. Thus, while the tripartite factors were not identified in the present study to have direct influences on test performance, their role in predicting other academic outcomes cannot be ruled out.

8.5 Conclusions

Academic tests are one of the most significant sources of stress for high school students across New Zealand and other cultures, yet the interplay between students' test anxiety and general emotional well-being has not been explicitly

studied. This study examined the influence of the tripartite dimensions of general emotional distress on more specific symptoms of depression, anxiety, and test anxiety. The pathways in which these factors may predict test performance were also explored. The findings established that the broad construct of negative affect, or the tendency to feel upset, reliably predicted greater levels of test-related worries and tension, which largely represented the cognitive and affective components of test anxiety, respectively. The findings suggest that negative affect is a useful point of focus for therapeutic interventions to improve student's emotional well-being. However, the broad construct of negative affect did not directly predict test performance, and its indirect effects could not be definitively established nor ruled out. However, the structural relationships among negative affect, test anxiety, and test performance illustrate the different trajectories through which general emotional distress may indirectly improve and impair test performance. The results highlight how the relationship between general negative emotions and test performance may only be understood in the context of other mediating factors, such as the presence of cognitive worries and apprehensive tension.

References

Achenbach, T. M., & Rescorla, L. A. (2001). *Manual for the ASEBA school-age forms and profiles.* Burlington: University of Vermont.

Allen, G., Elias, M., & Zlotlow, S. (1980). Behavioral interventions for alleviating test anxiety: A methodological overview of current therapeutic practices. In I. G. Sarason (Ed.), *Test anxiety: Theory, research, and applications* (pp. 155-185). Hillsdale, NJ: Erlbaum.

Alpert, R., & Haber, R. N. (1960). Anxiety in academic achievement situations. *The Journal of Abnormal and Social Psychology, 61,* 207-215. doi: 10.1037/h0045464

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

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed. text revision). Washington, DC: Author.

Anderson, E. R., & Hope, D. A. (2008). A review of the tripartite model for understanding the link between anxiety and depression in youth. *Clinical Psychology Review, 28*, 275-287. doi: 10.1016/j.cpr.2007.05.004

Anderson, E. R., Veed, G. J., Inderbitzen, H. M., & Hansen, D. J. (2010). An evaluation of the applicability of the tripartite constructs to social anxiety in adolescents. *Journal of Clinical Child & Adolescent Psychology, 39,* 195-207. doi: 10.1080/15374410903532643

Anderson, S. B., & Sauser, W. I. (1995). Measurement of test anxiety: An overview. In C. D. Spielberger & P. R. Vagg (Eds.), *Test anxiety: Theory, assessment, and treatment.* Washington, D.C.: Taylor & Francis.

Arbuckle, J. L. (2009). AMOS 18.0 [Computer software]. Chicago: SmallWaters.

Aronen, E. T., Vuontela, V., Steenari, M.-R., Salmi, J., & Carlson, S. (2005). Working memory, psychiatric symptoms, and academic performance at school. *Neurobiology of Learning and Memory, 83,* 33-42. doi: 10.1016/j.nlm.2004.06.010

Baddeley, A. D. (1996). Exploring the central executive. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 49A,* 5–28. doi: 10.1080/713755608

Baddeley, A. D. (2001). Is working memory still working? *American Psychologist, 56,* 851–864. doi: 10.1037/0003-066X.56.11.851

Bandalos, D. L., Yates, K., & Thorndike-Christ, T. (1995). Effects of math self-concept, perceived self-efficacy, and attributions for failure and success on

test anxiety. *Journal of Educational Psychology, 87,* 611-623. doi: 10.1037/0022-0663.87.4.611

Bandura, A., Pastorelli, C., Barbaranelli, C., & Caprara, G. V. (1999). Self-efficacy pathways to childhood depression. *Journal of Personality and Social Psychology, 76,* 258-269. doi: 10.1037/0022-3514.76.2.258

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51,* 1173-1182. doi: 10.1037/0022-3514.51.6.1173

Beck, R., Benedict, B., & Winkler, A. (2003). Depression and anxiety: Integrating the tripartite and cognitive content-specificity assessment models. *Journal of Psychopathology and Behavioral Assessment, 25,* 251-256. doi: 10.1023/A:1025899029040

Beidel, D. C., Turner, M. W., & Trager, K. N. (1994). Test anxiety and childhood anxiety disorders in African American and white school children. *Journal of Anxiety Disorders, 8*, 169-179. doi: 10.1016/0887-6185(94)90014-0

Benson, J. (1998). Developing a strong program of construct validation: A test anxiety example. *Educational Measurement: Issues and Practice, 17*, 10-17. doi: 10.1111/j.1745-3992.1998.tb00616.x

Benson, J., & Bandalos, D. L. (1992). Second-order confirmatory factor analysis of the Reactions to Tests scale with cross-validation. *Multivariate Behavioral Research, 27,* 459-487. doi: 10.1207/s15327906mbr2703_8

Benson, J., & El-Zahhar, N. (1994). Further refinement and validation of the Revised Test Anxiety Scale. *Structural Equation Modeling, 1*, 203-221.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107,* 238-246. doi: 10.1037/0033-2909.107.2.238

Birmaher, B., Khetarpal, S., Brent, d., Cully, M., Balach, L., Kaufman, J., & McKenzie Neer, S. (1997). The Screen for Child Anxiety Related emotional Disorders (SCARED): Scale construction and psychometric characteristics. *Journal of the American Academy of Child and Adolescent Psychiatry, 36*, 545-553. doi: 10.1097/00004583-199704000-00018

Blankstein, K. R., Flett, G. L., Boase, P., & Toner, B. B. (1990). Thought listing and endorsement measures of self-referential thinking in test anxiety. *Anxiety Research, 2,* 103-112. doi: 10.1080/08917779008249329

Blankstein, K. R., Toner, B. B., & Flett, G. L. (1989). Test anxiety and the contents of consciousness: Thought-listing and endorsement measures. *Journal of Research in Personality, 23*, 269-286. doi: 10.1016/0092-6566(89)9001-9

Blatz, W. E. (1925). The cardiac, respiratory, and electrical phenomena involve in emotion of fear. *Journal of Experimental Psychology, 8,* 109-132. doi: 10.1037/h0071039

Bradley, R., McCraty, R., Atkinson, M., Tomasino, D., Daugherty, D., & Arguelles, L. (2010). Emotion self-regulation, psychophysiological coherence, and test anxiety: Results from an experiment using electrophysiological measures. *Applied Psychophysiology and Biofeedback, 35,* 261-283. doi: 10.1007/s10484-010-9134-x

Brady, E. U., & Kendall, P. C. (1992). Comorbidity of anxiety and depression in children and adolescents. *Psychological Bulletin, 111*, 244-255. doi: 10.1037/0033-2909.111.2.244

Bray, M., & Kwok, P. (2003). Demand for private supplementary tutoring: Conceptual considerations, and socio-economic patterns in Hong Kong. *Economics of Education Review, 22*, 611-620. doi: 10.1016/S0272-7757(03)00032-3

Brown, C., Schulberg, H., Madonia, M., Shear, M., & Houck, P. (1996). Treatment outcomes for primary care patients with major depression and lifetime anxiety disorders. *American Journal of Psychiatry, 153,* 1293-1300.

Brown, C. H. (1938). Emotional reactions before examinations: II. Results of a questionnaire. *Journal of Psychology, 5,* 11-26. doi: 10.1080/00223980.1938.9917549

Brown, T. A., Chorpita, B. F., & Barlow, D. H. (1998). Structural relationships among dimensions of the DSM-IV anxiety and mood disorders and dimensions of negative affect, positive affect, and autonomic arousal. *Journal of Abnormal Psychology, 107,* 179-192. doi: 10.1037/0021-843X.107.2.179

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen, & J. S. Long (Eds.), *Testing structural equation models* (pp. 136-162). Newbury Park, CA: Sage.

Bull, R., Espy, K. A., & Wiebe, S. A. (2008). Short-term memory, working memory, and executive functioning in preschoolers: Longitudinal predictors of mathematical achievement at age 7 years. *Developmental Neuropsychology, 33,* 205-228. doi: 10.1080/87565640801982312

Byrne, B.M. (1989). *A primer of LISREL: Basic applications and programming for confirmatory factor analytic models*. New York: Springer-Verlag.

Byrne, B. M. (2010). *Structural equation modeling with AMOS: Basic concepts, applications, and programming* (2nd ed.). New York: Routledge.

Carmines, E.G., & McIver, J.P. (1981). Analyzing models with unobserved variables. In G. W. Bohrnstedt & E. F. Borgatta (Eds.), *Social measurement: Current issues*. Beverly Hills: Sage Publications.

Carr, A. (2006). *The handbook of child and adolescent clinical psychology: A contextual approach* (2nd ed.). New York: Routledge.

Carver, C. S., & Connor-Smith, J. (2010). Personality and coping. *Annual Review of Psychology, 61,* 679-704. doi: 10.1146/annurev.psych.093008.100352

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

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

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,* 268-274. doi: 10.1037/0022-0663.97.2.268

Chemers, M. M., Hu, L-T., & Garcia, B. F. (2001). Academic self-efficacy and first year college student performance and adjustment. *Journal of Educational Psychology, 93*, 55-64. doi: 10.1037/0022-0663.93.1.55

Chorpita, B. F. (2002). The tripartite model and dimensions of anxiety and depression: An examination of structure in a large school sample. *Journal of Abnormal Child Psychology*, *30*, 177–190. doi: 10.1023/A:1014709417132

Chorpita, B. F., Albano, A. M., & Barlow, D. H. (1998). The structure of negative emotions in a clinical sample of children and adolescents. *Journal of Abnormal Psychology, 107,* 74-85. doi: 10.1037/0021-843X.107.1.74

Chorpita, B. F., & Barlow, D. H. (1998). The development of anxiety: The role of control in the early environment. *Psychological Bulletin, 124*, 3-21. doi: 10.1037/0033-2909.124.1.3

Chorpita, B. F., & Daleiden, E. L. (2002). Tripartite dimensions of emotion in a child clinical sample: Measurement strategies and implications for clinical utility. *Journal of Consulting and Clinical Psychology, 70,* 1150-1160. doi: 10.1037/0022-006X.70.5.1150

Chorpita, B. F., Daleiden, E. L., Moffitt, C., Yim, L., & Umemoto, L. (2000). Assessment of the tripartite factors of emotion in children and adolescents I: Structural validity and normative data of an Affect and Arousal Scale. *Journal of Psychopathology and Behavioral Assessment, 22*, 141-160. doi: 10.1023/A:1007584423617

Chorpita, B. F., Plummer, C. M., & Moffitt, C. E. (2000). Relations of tripartite dimensions of emotion to childhood anxiety and mood disorders. *Journal of Abnormal Child Psychology, 28,* 299-310. doi: 10.1023/A:1005152505888

Clark, L. A. & Watson, D. (1991). Tripartite model of anxiety and depression: Psychometric evidence and taxonomic implications. *Journal of Abnormal Psychology, 100,* 316-336. doi: 10.1037/0021-843X.100.3.316

Clark, L. A., Watson, D., & Reynolds, S. (1995). Diagnosis and classification of psychopathology: Challenges to the current system and future directions. *Annual Review of Psychology, 46,* 121-153.

Cole, D. A., Truglio, R., & Peeke, L. (1997). Relation between symptoms of anxiety and depression in children: A multitrait-multimethod-multigroup assessment. *Journal of Consulting and Clinical Psychology, 65,* 110-119. doi: 10.1037/0022-006X.65.1.110

Cook, J. M., Orvaschel, H., Simco, E., Hersen, M., & Joiner, T. (2004). A test of the tripartite model of depression and anxiety in older adult psychiatric outpatients. *Psychology and Aging, 19,* 444-451. doi: 10.1037/0882-7974.19.3.444

Crawford, J. R., & Henry, J. D. (2004). The Positive and Negative Affect Schedule (PANAS): Construct validity, measurement properties and normative data in a large non-clinical sample. *British Journal of Clinical Psychology, 43,* 245-265. doi: 10.1348/0144665031752934

Curcio, G., Ferrara, M., & de Gennaro, L. (2006). Sleep loss, learning capacity and academic performance. *Sleep Medicine Reviews, 10*, 323-337. doi: 10.1016/j.smrv.2005.11.001

Daleiden, E., Chorpita, B., & Lu, W. (2000). Assessment of tripartite factors of emotion in children and adolescents II: Concurrent validity of the Affect and Arousal Scales for children. *Journal of Psychopathology and Behavioral Assessment, 22*, 161-182. doi: 10.1023/A:1007536507687

Daniels, L. M., Stupnisky, R. H., Pekrun, R., Haynes, T. L., Perry, R. P., & Newall, N. E. (2009). A longitudinal analysis of achievement goals: From affective antecedents to emotional effects and achievement outcomes. *Journal of Educational Psychology, 101,* 948-963. doi: 10.1037/a0016096

De Bolle, M., Decuyper, M., de Clercq, B., & de Fruyt, F. (2010). Relevance of the tripartite dimensions of affect for anxiety and depression in youth: Examining sex and psychopathology status. *Journal of Abnormal Child Psychology, 38,* 935-948. doi: 10.1007/s10802-010-9413-2

De Bolle, M. & de Fruyt, F. (2010). The tripartite model in childhood and adolescence: Future directions for developmental research. *Child Development Perspective, 4*, 174-180. doi: 10.1111/j.1750-8606.2010.00136.x

Deffenbacher, J. L. (1980). Worry and emotionality in test anxiety. In I. G. Sarason (Ed.), *Test anxiety: Theory, research, and applications* (pp. 111–124). Hillsdale, NJ: Erlbaum.

Derakshan, N., & Eysenck, M. W. (2009). Anxiety, processing efficiency, and cognitive performance: New developments from attentional control theory. *European Psychologist, 14,* 168-176. doi: 10.1027/1016-9040.14.2.168

Devine, A., Fawcett, K., Szűcs, D., & Dowker, A. (2012). Gender differences in mathematics anxiety and the relation to mathematics performance while controlling for test anxiety. *Behavioral and Brain Functions, 8* (33), 2-9.

Dyck, M. J., Jolly, J. B., & Kramer, T. (1994). An evaluation of positive affectivity, negative affectivity, and hyperarousal as markers for assessing between syndrome relationships. *Personality and Individual Differences, 17,* 637-646. doi: 10.1016/0191-8869(94)90136-8

Eley, T. C., & Stevenson, J. (2000). Specific life events and chronic experiences differentially associated with depression and anxiety in young twins. *Journal of Abnormal Child Psychology, 28,* 383-394. doi: 10.1023/A:1005173127117

Ellard, K. K., Fairholme, C. P., Boisseau, C. L., Farchione, T. J., & Barlow, D. H. (2010). Unified protocol for the transdiagnostic treatment of emotional disorders: Protocol development and initial outcome data. *Cognitive and Behavioral Practice, 17,* 88-101. doi: 10.1016/j.cbpra.2009.06.002

Elliot, A. J., McGregor, H. A., & Gable, S. (1999). Achievement goals, study strategies, and exam performance: A mediational analysis. *Journal of Educational Psychology, 91*, 549-563. doi: 10.1037/0022-0663.91.3.549

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

Eysenck, M. W., & Calvo, M. G. (1992). Anxiety and performance: The processing efficiency theory. *Cognition and Emotion, 6*, 409–434. doi: 10.1080/02699939208409696

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

Feldman, L. A. (1993). Distinguishing depression and anxiety in self-report: Evidence from confirmatory factor analysis on nonclinical and clinical sample. *Journal of Consulting and Clinical Psychology, 61,* 631-638. doi: 10.1037/0022-006X.61.4.631

First, M. B. (2007). *Comorbidity of depression and generalized anxiety disorder.* Retrieved from www.dsm5.org/research/pages/comorbidityofdepressionandgeneralizedanxietydisorder(june20-22,2007).aspx

Fletcher, C., Lovatt, C., & Baldry, C. (1997). A study of state, trait, and test anxiety, and their relationship to assessment center performance. *Journal of Social Behavior and Personality, 12,* 205-214.

Flett, G. L., Blankstein, K. R., & Boase, P. (1987). Self-focused attention in test anxiety and depression. *Journal of Social Behavior and Personality, 2*, 259-266.

Forsyth, D. R., Story, P. A., Kelley, K. N., & McMillan, J. H. (2009). What causes failure and success? Students' perceptions of their academic outcomes. *Social Psychology of Education, 12,* 157-174. doi: 10.1007/s11218-008-9078-7

Fox-Wasylyshyn, S. M. & El-Masri, M. M. (2005). Handling missing data in self-report measures. *Research in Nursing & Health, 28,* 488-495. doi: 10.1002/nur.20100

Fröjd, S. A., Nissinen, E. S., Pelkonen, M. U., Marttunen, M. J., Koivisto, A. M., & Kaltiala-Heino, R. (2008). Depression and school performance in middle adolescent boys and girls. *Journal of adolescence*, *31*, 485-498. doi: 10.1016/j.adolescence.2007.08.006

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

Fritz, M. S., & MacKinnon, D. P. (2007). Required sample size to detect the mediated effect. *Psychological Science, 18,* 233-239. doi: 10.1111/j.1467-9280.2007.01882.x

Garber, J., & Horowitz, J. L. (2002). Depression in children. In I. H. Gotlib, & C. L. Hammen (Eds.), *Handbook of depression* (pp. 510-540). New York: Guilford Press.

Gregor, A. (2005). Examination anxiety: Live with it, control it or make it work for you? *School Psychology International, 26*, 617-635. doi: 10.1177/0143034305060802

Haines, M. E., Norris, M. P., & Kashy, D. A. (1996). The effects of depressed mood on academic performance in college students. *Journal of College Student Development, 37*, 519-526.

Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2006). *Multivariate data analysis* (6th ed.). New Jersey: Pearson Prentice Hall.

Harvard School of Medicine. (2005). *National comorbidity survey (NCS) and national comorbidity survey replication (NCS-R).* Retrieved from www.hcp.med.harvard.edu/ncs

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

Hettema, J. M. (2008). What is the genetic relationship between anxiety and depression? *American Journal of Medical Genetics Part C: Seminars in Medical Genetics, 148C,* 140-146. doi: 10.1002/ajmg.c.30171

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

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,* 1-55. doi: 10.1080/10705519909540118

Hughes, A. A., & Kendall, P. C. (2009). Psychometric properties of the Positive and Negative Affect Scale for Children (PANAS-C) in children with anxiety disorders. *Child Psychiatry and Human Development, 40,* 343-352. doi: 10.1007/s10578-009-0130-4

Hughes, A. A., Lourea-Waddell, B., & Kendall, P. C. (2008). Somatic complaints in children with anxiety disorders and their unique prediction of poorer academic performance. *Child Psychiatry and Human Development, 39,* 211-220. doi: 10.1007/s10578-007-0082-5

Humphreys, M. S., & Revelle, W. (1984). Personality, motivation, and performance: A theory of the relationship between individual differences and information processing. *Psychological Review, 91,* 153-184. doi: 10.1037/0033-295X.91.2.153

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

Hysenbegasi, A., Hass, S. L., & Rowland, C. R. (2005). The impact of depression on the academic productivity of university students. *The Journal of Mental Health Policy and Economics, 8,* 145-151.

Jacques, H. A., & Mash, E. J. (2004). A test of the tripartite model of anxiety and depression in elementary and high school boys and girls. *Journal of Abnormal Child Psychology, 32*, 13-25. doi: 10.1023/B:JACP.0000007577.38802.18

Joiner, T. E. (1996). A confirmatory factor-analytic investigation of the tripartite model of depression and anxiety in college students. *Cognitive Therapy and Research, 20,* 521-539. doi: 10.1007/BF02227911

Joiner, T. E., Catanzaro, S. J., & Laurent, J. (1996). Tripartite structure of positive and negative affect, depression, and anxiety in child and adolescent psychiatric inpatients. *Journal of Abnormal Psychology, 105*, 401-409. doi: 10.1037/0021-843X.105.3.401

Joiner, T. E., & Lonigan, C. J. (2000). Tripartite model of depression and anxiety in youth psychiatric inpatients: Relations with diagnostic status and future symptoms. *Journal of Clinical Child Psychology, 29,* 372-382. doi: 10.1207/S15374424JCCP2903_8

Joiner, T. E., Steer, R. A., Beck, A. T., Schmidt, N. B., Rudd, M. D., & Catanzaro, S. J. (1999). Physiological hyperarousal: Construct validity of a central aspect of the tripartite model of depression and anxiety. *Journal of Abnormal Psychology, 108*, 290-298. doi: 10.1037/0021-843X.108.2.290

Kauer-Sant'Anna, M., Frey, B., Andreazza, A., Cereser, K., Gazalle, F., Tramontina, J., … Kapczinski, F. (2007). Anxiety comorbidity and quality of life in bipolar disorder patients. *Canadian Journal of Psychiatry, 52,* 175-181.

Keogh, E., Bond, F. W., French, C. C., Richards, A., & Davis, R. E. (2004). Test anxiety, susceptibility to distraction and examination performance. *Anxiety, Stress, and Coping, 17,* 241-252. doi: 10.1080/10615300410001703472

Kessler, R. C., & Walters, E. E. (1998). Epidemiology of DSM-III-R major depression and minor depression among adolescents and young adults in the national comorbidity survey. *Depression and Anxiety, 7,* 3-14. doi: 10.1002/(SICI)1520-6394(1998)7:1<3::AID-DA2>3.0.CO;2-F

Kiernan, G., Laurent, J., Joiner, T. E., Catanzaro, S. J., & MacLachlan, M. (2004). Cross-cultural examination of the tripartite model with children: Data from the Barretstown Studies. *Journal of Personality Assessment, 77*, 359-379. doi: 10.1207/S15327752JPA7702_15

King, N., Mietz, A., Tinney, L., & Ollendick, T. (1995). Psychopathology and cognition in adolescents experience severe test anxiety. *Journal of Clinical Child Psychology, 24*, 49-54. doi: 10.1207/s15374424jccp2401_6

Kirkland, K., & Hollandsworth, J. G. (1979). Test anxiety, study skills, and academic performance. *Journal of College Student Personnel, 20,* 431-435.

Kirkland, K., & Hollandsworth, J. G. (1980). Effective test taking: Skills-acquisition versus anxiety-reduction techniques. *Journal of Consulting and Clinical Psychology, 48,* 431-439. doi: 10.1037/0022-006X.48.4.431

Kleijn, W. C., van der Ploeg, H. M., & Topman, R. M. (1994). Cognition, study habits, test anxiety, and academic performance. *Psychological Reports, 75,* 1219-1226.

Kring, A. M., Person, J. B., & Thomas, C. (2007). Changes in affect during treatment for depression and anxiety. *Behaviour Research and Therapy, 45,* 1753-1764. doi: 10.1016/j.brat.2007.02.001

Kouzma, N. M., & Kennedy, G. A. (2004). Self-reported sources of stress in senior high school students. *Psychological Reports, 94,* 314-316. doi: 10.2466/pr0.94.1.314-316

Kovacs, M. (1981). Rating scales to assess depression in school-aged children. *Acta Paedopsychiatrica: International Journal of Child & Adolescent Psychiatry, 46,* 305-315.

Kwok, P. (2004). Examination-oriented knowledge and value transformation in East Asian cram schools. *Asia Pacific Education Review, 5*, 64-75. doi: 10.1007/BF03026280

Landis, C. (1924). Studies of emotional reactions: I. A preliminary study of facial expression. *Journal of Experimental Psychology, 7,* 325-341. doi: 10.1037/h0076072

Latzman, R. D., Elkovitch, N., Young, J., & Clark, L. A. (2010). The contribution of executive functioning to academic achievement among male adolescents. *Journal of clinical and experimental neuropsychology, 32*, 455-462. doi: 10.1080/13803390903164363

Laurent, J., Catanzaro, S., Joiner, T., Rudolph, K., Potter, K., Lambert, S., … Gathright, T. (1999). A measure of positive and negative affect for children: Scale development and preliminary validation. *Psychological Assessment, 11,* 326-338. doi: 10.1037/1040-3590.11.3.326

Laurent, J., Joiner, T. E., & Catanzaro, S. J. (2011). Positive affect, negative affect, and physiological hyperarousal among referred and nonreferred youths. *Psychological Assessment, 23,* 945-957. doi: 10.1037/a0024080

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,* 195-208. doi: 10.1002/per.2410030305

Lazaratou, H., Dikeos, D. G., Anagnostopoulos, D. C., & Soldatos, C. R. (2010). Depressive symptomatology in high school students: The role of age, gender and academic pressure. *Community Mental Health Journal*, *46*, 289-295. doi:10.1007/s10597-010-9296-2

Lewinsohn, P. M., & Essau, C. A. (2002). Depression in adolescents. In I. H. Gotlib, & C. L. Hammen (Eds.), *Handbook of depression* (pp. 541-559). New York: Guilford Press.

Liebert, R. M., & Morris, L. W. (1967). Cognitive and emotional components of test anxiety: A distinction and some initial data. *Psychological Reports, 20*, 975–978. doi: 10.2466/pr0.1967.20.3.975

Little, T. D., Cunningham, W. A., Shahar, G., & Widaman, K. F. (2002). To parcel or not to parcel : Exploring the question, weighing the merits. *Structural Equation Modeling : A Multidisciplinary Journal, 9,* 151-173. doi : 10.1207/S15328007SEM0902_1

Lonigan, C. J., Carey, M. P., & Finch, A. J. (1994). Anxiety and depression in children and adolescents: Negative affectivity and the utility of self-reports. *Journal of Consulting and Clinical Psychology, 62,* 1000-1008. doi: 10.1037/0022-006X.62.5.1000

Lonigan, C. J., Hooe, E. S., David, C., & Kistner, J. (1999). Positive and negative affectivity in children: Confirmatory factor analysis of a two-factor model and its relation to symptoms of anxiety and depression. *Journal of Consulting and Clinical Psychology, 67,* 374-386. doi: 10.1037/0022-006X.67.3.374

Lonigan, C. J., Phillips, B. M., & Hooe, E. S. (2003). Relations of positive and negative affectivity to anxiety and depression in children: Evidence from a latent variable longitudinal study. *Journal of Consulting and Clinical Psychology, 71*, 465-481. doi: 10.1037/0022-006X.71.3.465

Lovibond, S. H. & Lovibond, P. F. (1993). *Manual for the Depression Anxiety Stress Scales (DASS).* Psychology Foundation of Australia.

Lowe, P. A., Lee, S.W.,Witteborg, K. M., Pritchard, K.W., Luhr, M. E., Cullinan, C. M., … Janik, M. (2008). The Test Anxiety Inventory for Children and Adolescents (TAICA): Examination of the psychometric properties of a new multidimensional measure of test anxiety among elementary and secondary school students. *Journal of Psychoeducational Assessment, 26*, 215-230. doi: 10.1177/0734282907303760

Ma, X., & Xu, J. (2004). Determining the causal order between attitude toward mathematics and achievement in mathematics. *American Journal fof Education, 110,* 256-280.

MacKinnon, D. P., Lockwood, C. M., Hoffman, J. M., West, S. G., & Sheets, V. (2002). A comparison of methods to test mediation and other intervening variable effects. *Psychological Methods, 7,* 83-104. doi:10.1037/1082-989X.7.1.83

Mandler, G., & Sarason, S. B. (1952). A study of anxiety and learning. *The Journal of Abnormal and Social Psychology, 47,* 166-173. doi: 10.1037/h0062855

March, J. S. (1997). *Multidimensional Anxiety Scale for Children: Technical manual*. North Tonawanda, NY: Multi-Health Systems.

Martin, A., & Marsh, H. (2003). Fear of failure: Friend of foe*? Australian Psychologist, 38,* 31-38. doi: 10.1080/00050060310001706997

Massey University. (n.d.). *Admission to undergraduate study.* Retrieved from http://www.massey.ac .nz/massey/admission/entry-requirements/nz_citizens_permanent_residents/undergraduate.cfm

McDonald, A. (2001). The prevalence and effects of test anxiety in school children. *Educational Psychology, 21,* 89-101. doi: 10.1080/01443410020019867

Meijer, J. (2001). Learning potential and anxious tendency: Test anxiety as a bias factor in educational testing. *Anxiety, Stress & Coping, 14*, 337-362. doi: 10.1080/10615800108248361

Milfont, T. L., Merry, S., Robinson, E., Denny, S., Crengle, S., & Ameratunga, S. (2008). Evaluating the short form of the Reynolds Adolescent Depression Scale in New Zealand adolescents. *Australian and New Zealand Journal of Psychiatry, 42*, 950-954. doi: 10.1080/00048670802415343

Miller, H., & Bichsel, J. (2004). Anxiety, working memory, gender, and math performance. *Personality* and Individual Differences, 37, 591-606. doi: 10.1016/j.paid.2003.09.029

Mineka, S., Watson, D., & Clark, L. (1998). Comorbidity of anxiety and unipolar mood disorders. *Annual Review of Psychology, 49,* 337-403.

Moffitt, T. E., Caspi, A., Harrington, H. L., Milne, B. J., Melchior, M., Goldberg, D., & Poulton, R. (2007). Generalized anxiety disorder and depression: Childhood risk factors in a birth cohort followed to age 32. *Psychological Medicine, 27*, 441-452. doi: 10.1017/S0033291706009640

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,* 541-555. doi: 10.1037/0022-0663.73.4.541

Muris, P. (1997). *The Screen for Child Anxiety Related Emotional Disorders (revised version)*. Maastricht: Maastricht University.

Muris, P., Dreessen, L., Bögels, S., Weckx, M., & Melick, M. (2004). A questionnaire for screening a broad range of DSM-defined anxiety disorder symptoms in clinically referred children and adolescents. *Journal of Child Psychology and Psychiatry, 45*, 813-820. doi: 10.1111/j.1469-7610.2004.00274.x

Muris, P., Merckelbach, H., Ollendick, T., King, N., & Bogie, N. (2002). Three traditional and three new childhood anxiety questionnaires: Their reliability and validity in a normal adolescent sample. Behaviour Research and Therapy, 40, 753-772. doi: 10.1016/S0005-7967(01)00056-0

Muris, P., Merckelbach, H., Schmidt, H., & Mayer, B. (1999). The revised version of the Screen for Child Anxiety Related Emotional Disorders (SCARED-R): Factor structure in normal children. *Personality and Individual Differences, 26*, 99-112. doi: 10.1016/S0191-8869(98)00130-5

Muris, P., Merckelbach, H., van Brakel, A., & Mayer, B. (1999). The revised version of the Screen for Child Anxiety Related Emotional Disorders (SCARED-R): Further evidence for its reliability and validity. *Anxiety, Stress, and Coping, 12,* 411-425. doi: 10.1080/10615809908249319

Musch, J., & Bröder, A. (1999). Test anxiety versus academic skills: A comparison of two alternative models for predicting performance in a statistics exam. *British Journal of Educational Psychology, 69,* 105-116. doi: 10.1348/000709999157608

Myers, K., & Winters, N. (2002). Ten-year review of rating scales: II. Scales for internalizing disorders. *Journal of American Academy of Child & Adolescent Psychiatry, 41,* 634-659. doi: 10.1097/00004583-200206000-00004

Naragon-Gainey, K., Watson, D., & Markon, K. E. (2009). Differential relations of depression and social anxiety symptoms to the facets of extraversion/positive emotionality. *Journal of Abnormal Psychology, 118,* 299-310. doi: 10/1037/a0015637

Nath, S. R. (2008). Private supplementary tutoring among primary students in Bangladesh. *Educational Studies, 34*, 55-72. doi: 10.1080/03055690701785285

Naveh-Benjamin, M., McKeachie, W. J., & Lin, Y.-G. (1987). Two types of test-anxious students: Support for an information processing model. *Journal of Educational Psychology, 79,* 131-136. doi: 10.1037/0022-0663.79.2.131

Naveh-Benjamin, M., McKeachie, W. J., Lin, Y.-G., & Holinger, D. P. (1981). Test anxiety: Deficits in information processing. *Journal of Educational Psychology, 73,* 816-824. doi: 10.1037/0022-0663.73.6.816

New Zealand Qualifications Authority. (n.d.). *Credits and the NCEA.* Retrieved from http://www.nzqa.govt.nz/assets/qualifications-and-standards/qualifications/ncea/Understanding-NCEA/factsheet-4.pdf

Nicholson, W. M. (1958). The influence of anxiety upon learning: Interference or drive increment? *Journal of Personality, 26*, 303-319. doi: 10.1111/j.1467-6494.1958.tb01587.x

Oakley Browne, M. A., Wells, E., Scott, K., & McGee, M. A. (2006). Lifetime prevalence and projected lifetime risk of DSM-IV disorders in Te Rau Hinengaro: The New Zealand Mental Health Survey. *Australian and New Zealand Journal of Psychiatry, 40,* 865-874. doi: 10.1111/j.1440-1614.2006.01905.x

Ollendick, T. H., Seligman, L. D., Goza, A. B., Byrd, D. A., & Singh, K. (2003). Anxiety and depression in children and adolescents: A factor-analytic examination of the tripartite model. *Journal of Child and Family Studies, 12,* 157-170. doi: 10.1023/A:1022806731527

Otto, M., Powers, M., Stathopoulou, G., & Hofmann, S. (2008). Panic disorder and social phobia. In M. Whisman (Ed.), *Adapting cognitive therapy for depression: Managing complexity and comorbidity* (pp. 185-208). New York: The Guildford Press.

Owens, M., Stevenson, J., Hadwin, J., & Norgate, R. (2012). Anxiety and depression in academic performance: An exploration of the mediating factors of worry and working memory. *School Psychology International, 33*, 433-449. doi: 10.1177/0143034311427433

Owens, M., Stevenson, J., Norgate, R., & Hadwin. J. (2008). Processing efficiency theory in children: Working memory as a mediator between trait anxiety and academic performance. *Anxiety, Stress, & Coping, 21*, 417-430. doi: 10.1080/10615800701847823

Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review, 18,* 315-341. doi: 10.1007/s10648-006-9029-9

Pekrun, R., Frenzel, A. C., Goetz, T., & Perry, R. P. (2007). The control-value theory of achievement emotions: An integrative approach to emotions in education. In P. Schutz & R. Pekrun (Eds.), *Emotion in education* (pp. 13-36

Amsterdam: Academic Press.

Pekrun, R., Goetz, T., Perry, R., Kramer, L., Hochstadt, M., & Molfenter, S. (2004). Beyond test anxiety: Development and validation of the test emotions questionnaire (TEQ). *Anxiety, Stress, and Coping, 17,* 287-316. doi: 10.1080/10615800412331303847

Perkins, A. M., Kemp, S. E., & Corr, P. J. (2007). Fear and anxiety as separable emotions: An investigation of the revised reinforcement sensitivity theory of personality. *Emotion, 7*, 252-261. doi: 10.1037/1528-3542.7.2.252

Piko, B. (2001). Gender differences and similarities in adolescents' ways of coping. *The Psychological Record, 51*, 223-235.

Pintrich, P. R., & de Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology, 82*, 33-40. doi: 10.1037/0022-0663.82.1.33

Reich, J., Warshaw, M., Peterson, L., & White, K. (1993). Comorbidity of panic and major depressive disorder. *Journal of Psychiatric Research, 27* (Supplement), 23-33. doi: 10.1016/0022-3956(93)90015-T

Reynolds, C. R., & Richmond, B. O. (1985). *Manual for the Revised Children's Manifest Anxiety Scale.* Los Angeles: Western Psychological Services.

Reynolds, W. M. (1986). *Reynolds Adolescent Depression Scale.* John Wiley & Sons, Inc.

Reynolds, W. M. (2002). *Reynolds Adolescent Depression Scale – Second Edition: Professional Manual.* Odessa, FL: Psychological Assessment Resources.

Rocklin, T., & Ren-Min, Y. (1989). Development and adaptation of the Chinese Test Anxiety Inventory: A research note. In R. Schwarzer, H. M. van der Ploeg, & C. D. Spielberger (Eds.), *Advances in test anxiety research* (Vol. 6, pp. 245–251). Lisse, The Netherlands: Swets & Zeitlinger.

Roth, P.L., Switzer, F. S., & Switzer, D. M. (1999). Missing data in multiple item scales: A monte carlo analysis of missing data techniques. *Organizational Research Methods, 2*, 211-232. doi: 10.1177/109442819923001

Ruini, C., Ottolini, F., Tomba, E., Belaise, C., Albieri, E., Visani, D., … Fava, G. A. (2009). School intervention for promoting psychological well-being in adolescence. *Journal of Behavior Therapy and Experimental Psychiatry, 40,* 522-532. doi: 10.1016/j.jbtep.2009.07.002

Sadock, B. J., & Sadock, V. A. (2007). *Kaplan and Sadock's synopsis of psychiatry: Behavioral Science / Clinical Psychiatry* (10th ed.). Philadephia: Wolter Kluwer / Lippincott Williams & Wilkins.

Sarason, I. G. (1984). Stress, anxiety, and cognitive interference: Reactions to Tests. *Journal of Personality and Social Psychology, 46***,** 929–938. doi: 10.1037/0022-3514.46.4.929

Sarason, I. G. (1986). Test anxiety, worry, and cognitive interference. In R. Schwarzer (Ed.), *Self-related cognitions in anxiety and motivation* (pp. 19-34). Hillsdale, NJ: Erlbaum.

Sarason, S., & Mandler, G. (1952). Some correlates of test anxiety. *Journal of Abnormal and Social Psychology, 47,* 810-817. doi: 10.1037/h0060009

Savin, N. E., & White, K. J. (1977). The Durbin-Watson test for serial correlation with extreme sample sizes or many regressors. *Econometrica, 45*, 1989-1996.

Schumacker, R. E., & Lomax, R. G. (2004). *A beginner's guide to structural equation modeling* (2nd ed.). New Jersey: Lawrence Erlbaum Associates.

Scott, K., McGee, M. A., Oakley Browne, M., & Wells, E. (2006). Mental disorder comorbidity in Te Rau Hinengaro: The New Zealand Mental Health Survey. *Australian and New Zealand Journal of Psychiatry, 40,* 875-881. doi: 10.1111/j.1440-1614.2006.01906.x

Seipp, B. (1991). Anxiety and academic performance: A meta-analysis of findings. *Anxiety Research, 4,* 27-41. doi: 10.1080/08917779108248762

Seligman, L. D., & Ollendick, T. H. (1998). Comorbidity of anxiety and depression in children and adolescents: An integrative review. *Clinical Child and Family Psychology Review, 1,* 125-144. doi: 10.1023/A:1021887712873

Seligman, L. D., Ollendick, T. H., Langley, A. K., & Baldacci, H. B. (2004). The utility of measures of child and adolescent anxiety: A meta-analytic review of the Revised Children's Manifest Anxiety Scale, the State-Trait Anxiety Inventory for Children, and the Child Behavior Checklist. *Journal of Clinical Child & Adolescent Psychology, 33,* 557-565. doi: 10.1207/s15374424jccp3303_13

Sellbom, M., Ben-Porath, Y. S., & Bagby, R. M. (2008). On the hierarchical structure of mood and anxiety disorders: Confirmatory evidence and elaboration of a model of temperament markers. *Journal of Abnormal Psychology, 117,* 576-590. doi: 10.1037/a0012536

Sharma, S., & Sud, A. (1990). Examination stress and test anxiety: A cross-cultural perspective. *Psychology and Developing Societies, 2,* 183-201. doi: 10.1177/097133369000200203

Shulruf, B., Wang, Y. G., Zhao, Y. J., & Baker, H. (2011). Rethinking the admission criteria to nursing school. *Nurse Education Today, 31,* 727-732. doi: 10.1016/j.nedt.2010.11.024

Silverman, W. K., & Ollendick, T. H. (2005). Evidence-based assessment of anxiety and its disorders in children and adolescents. *Journal of Clinical Child and Adolescent Psychology*, *34*, 380-411. doi: 10.1207/s15374424jccp3403_2

Simon, N., Otto, M., Wisniewski, S., Fossey, M., Sagduyu K., Frank, E., … Pollack, M. (2004). Anxiety disorder comorbidity in bipolar disorder patients: Data from the first 500 participants in the systematic treatment enhancement program for bipolar disorder (STEP-BD). *The American Journal of Psychiatry, 161*, 2222-2229. doi: 10.1176/appi.ajp.161.12.2222

Spence, S. H. (1998). A measure of anxiety symptoms among children. *Behaviour Research and Therapy, 36,* 545–566. doi: 10.1016/S0005-7967(98)00034-5

Spicer, J. (2005). *Making sense of multivariate data analysis.* California: Sage Publications.

Spielberger, C. D. (1972). Anxiety as an emotional state. In C. D. Spielberger (Ed.), *Anxiety: Current trends in theory and research* (Vol. 1, pp. 23-49). New York: Academic Press.

Spielberger, C. D. (1973). *Manual for the State–Trait Anxiety Inventory for Children*. Palo Alto, CA: Consulting Psychologists Press.

Spielberger, C. D., Gonzalez, H., Taylor, C. J., Aglaze, B., & Anton, W. D. (1978). Examination Stress and test anxiety. In C.D. Spielberger, & I. Sarason (Eds.), *Stress and anxiety* (Vol.5, pp.167-191). New York: Hemisphere/Wiley.

Spielberger, C. D., & Vagg, P. R. (1995). Test anxiety: A transactional process model. In C. D. Spielberger, & P. R. Vagg (Eds.), *Test anxiety: Theory, assessment, and treatment* (pp. 3-14). Washington, D.C.: Taylor & Francis.

SPSS Inc. (2007). SPSS 16.0 for Windows [Computer software]. Chicago: Author.

Stark, K. D., & Laurent, J. (2001). Joint factor analysis of the Children's Depression Inventory and the Revised Children's Manifest Anxiety Scale. *Journal of Clinical Child & Adolescent Psychology, 30,* 552-567. doi: 10.1207/S15374424JCCP3004_11

Stevanovic, D., Laurent, J., & Lakic, A. (2013). Measuring positive and negative affect and physiological hyperarousal among Serbian youth. *Journal of Personality Assessment, 95,* 107-117. doi: 10.1080/00223891.2012.718301

Stevens, J. P. (2002). *Applied multivariate statistics for the social sciences* (4th ed.). Hillsdale, NJ: Lawrence Erlbaum.

Struthers, C. W., Perry, R. P., & Menec, V. H. (2000). An examination of the relationship among academic stress, coping, motivation, and performance in college. *Research in Higher Education, 41* (5), 581-592. doi: 10.1023/A:1007094931292

Tabachnick, B. G., & Fidell, L. S. (2000). *Using multivariate statistics* (4th ed.). Boston, M.A.: Allyn and Bacon.

Teachman, B. A., Siedlecki, K. L., & Magee, J. C. (2007). Aging and symptoms of anxiety and depression: Structural invariance of the tripartite model. *Psychology and Aging, 22*, 160-170. doi: 10.1037/0882-7974.22.1.160

Tellegen, A. (1985). Structures of mood and personality and their relevance to assessing anxiety, with an emphasis on self-report. In A. H. Tuma, & J. D. Maser (Eds.), *Anxiety and the anxiety disorders* (pp. 681-706). Hillsdale, N.J. England: Lawrence Erlbaum Associates.

Thelwell, R. C., Lane, A. M., & Weston, N. J. (2007). Mood states, self-set goals, self-efficacy and performance in academic examinations. *Personality and Individual Differences, 42,* 573-583. doi: 10.1016/j.paid.2006.07.024

Tobias, S. (1985). Test anxiety: Inference, defective skills, and cognitive capacity. *Educational Psychologist, 20,* 135-142.

Tuccitto, D. E., Giacobbi, P. R., & Leite, W. L. (2009). The internal structure of positive and negative affect: A confirmatory factor analysis of the PANAS. *Educational and Psychological Measurement, 70*, 125-141. doi: 10.1177/0013164409344522

Tully, P. J., Zajac, I. T., & Venning, A. J. (2009). The structure of anxiety and depression in a normative sample of younger and older Australian adolescents. *Journal of Abnormal Child Psychology, 37,* 717-726. doi: 10.1007/s10802-009-9306-4

Van Ameringen, M., Mancini, C., & Farvolden, P. (2003). The impact of anxiety disorders on educational achievement. *Journal of Anxiety Disorder, 17,* 561-571. doi: 10.1016/S0887-6185(02)00228-1

Vasey, M. W., El-Hag, N., & Daleiden, E. L. (1996). Anxiety and the processing of emotionally threatening stimuli: Distinctive patterns of selective attention among high- and low- test anxious children. *Child Development, 67,* 1173-1185. doi: 10.1111/j.1467-8624.1996.tb01789.x

Vieta, E., Colom, F., Martinez-Aran, A., Benabarre, A., Reinares, M., & Gasto, C. (2000). Bipolar II disorder and comorbidity. *Comprehensive Psychiatry, 41,* 339-343. doi: 10.1053/comp.2000.9011

Von der Embse, N., Barterian, J., & Segool, N. (2013). Test anxiety interventions for children and adolescents: A systematic review of treatment studies from 2000-2010. *Psychology in the Schools, 50,* 57-71. doi: 10.1002/pits.21660

Watson, D., Clark, L. A., & Carey, G. (1988). Positive and negative affectivity and their relation to anxiety and depressive disorders. *Journal of Abnormal Psychology, 97*, 346-353. doi: 10.1037/0021-843X.97.3.346

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology, 54,* 1063-1070. doi: 10.1037/0022-3514.54.6.1063

Watson, D., Clark, L. A., Weber, K., Assenheimer, J. S., Strauss, M. E., & McCormick, R. A. (1995). Testing a tripartite model: II. Exploring the symptom structure of anxiety and depression in student, adult, and patient samples. *Journal of Abnormal* Psychology, 104, 15-25. doi: 10.1037/0021-843X.104.1.15

Watson, D., Gamez, W., & Simms, L. (2005). Basic dimensions of temperament and their relation to anxiety and depression: A symptom-based perspective. *Journal of Research in Personality, 39*, 46-66. doi: 10.1016/j.jrp.2004.09.006

Wells, E., Oakely Browne, M., Scott, K., McGee, M., Baxter, J., & Kokaua, J. (2006). Te Rau Hinengaro: The New Zealand mental health survey: Overview of methods and findings. *Australian and New Zealand Journal of Psychiatry, 40,* 835-844. doi: 10.1111/j.1440-1614.2006.01902.x

West, S. G., Finch, J. F., & Curran, P. J. (1995). Structural equation models with nonnormal variables: Problems and remedies. In R. Hoyle (Ed.), *Structural equation modelling: Concept, issues, and applications* (pp. 56-75). London: Sage Publications.

Wheaton, B., Muthén, B., Alwin, D.F. & Summers, G.F. (1977). Assessing reliability and stability in panel models. In D. R. Heise (Ed.), *Sociological methodology,* (pp. 84-136). San Francisco: Jossey-Bass.

Williams, S. L. (2005). Anxiety disorders. In J. E. Maddux & B. A. Winstead (Eds.), *Psychopatholgy: Foundations for a contemporary understanding* (pp. 127-154). New Jersey: Lawrence Erlbaum Associates, Inc., Publishers.

Wine, J. D. (1971). Test anxiety and direction of attention. *Psychological Bulletin, 76,* 92-104. doi: 10.1037/h0031332

Wine. J. D. (1980). Cognitive-attentional theory of test anxiety. In I. Sarason (Ed.), *Test anxiety: Theory, research, and applications* (pp. 349-385). Hillsdale, N.J.: Erlbaum.

Wolpe, J. (1950). Need-reduction, drive-reduction, and reinforcement: A neurophysiological view. *Psychological Review, 57,* 19-26. doi: 10.1037/h0055810

Zeidner, M. (1998). *Test anxiety: The state of the art.* New York: Plenum Press.

Zimmer, J., Hocevar, D., Bachelor, P., & Meinke, D. L. (1992). An analysis of the Sarason (1984) four-factor conceptualization of test anxiety. *Advances in Test Anxiety Research*, *7*, 103-113.

Zohar, D. (1998). An additive model of test anxiety: Role of exam-specific expectations. *Journal of Educational Psychology, 90,* 330-340. doi: 10.1037/0022-0663.90.2.330

Zychinski, K., & Polo, A. J. (2012). Academic achievement and depressive symptoms in low-income Latino youth. *Journal of Child and Family Studies, 21,* 565-577. doi: 10.1007/s10826-011-9509-5

Appendices

Appendix A – Information Sheet

MASSEY UNIVERSITY
COLLEGE OF HUMANITIES
AND SOCIAL SCIENCES
TE KURA PŪKENGA TANGATA

New Zealand Survey of Students' Emotions

INFORMATION SHEET

Researcher's Introduction

My name is Edwin Chin, and I'm from the School of Psychology at Massey University. This study is part of my doctoral dissertation, and is towards a degree in Doctor of Clinical Psychology. The supervisors of this study are Dr. Mei-Wah Williams (Senior Lecturer), Dr. Shane Harvey (Director of Massey Psychology Clinic), and Dr. Joanne Taylor (Senior Lecturer and Clinician).

Purposes of the Research:

Studies have found that students who are highly anxious during exams perform less well compared to those who are less anxious. This research firstly aims to replicate this finding in New Zealand. Then, it aims to identify which types of moods are more strongly involved in producing test anxiety and impairing academic performance. This will help school counsellors, teachers, and psychologists be able to help students who struggle to perform at their best under anxiety provoking situations. It will also help you as students have a better understanding of the impact of moods on your studies.

Your Participation:

This study needs Year 12 and 13 high school students, aged 16 and above, doing NCEA English and Mathematics at any level (either Mathematics with Statistics and/or Calculus at level 3). You are invited to fill out a questionnaire about the experience of worry and anxiety over tests and exams. Questions will also ask about your mood and how you generally feel. This process will take around 30 minutes. In February 2011, I will send you an email asking for your permission to obtain your NCEA English and Mathematics achievement standards results for years 2009 and 2010.

Your participation is confidential, and any information you provide will not be shared with anyone including your peers, teachers, or parents. Neither grades nor academic relationships with your school and teachers will be affected whether or not you choose to participate.

Your questionnaire will be given a code number and kept separate from your name and contact information. Completed questionnaires will be stored in a locked filing cabinet, and will be destroyed at the end of the five-year storage period.

Your Rights

You are under no obligation to accept this invitation. If you decide to participate, you have the right to:

- decline to answer any particular question;
- withdraw from the study at any time during participation;
- ask any questions about the study at any time during participation;
- provide information on the understanding that your name will not be used unless you give permission to the researcher;
- be given access to a summary of the project findings when it has concluded.

Support services

The questionnaire contains items that describe symptoms of depression and other negative feelings. It is normal for anyone to experience some of these symptoms at any point in time, and scoring high on the questionnaire does not mean you are depressed. However, if you experience distress or depression, it is recommended that you seek out available support services such as your school counsellor. There are also free telephone services available, such as the Depression Helpline: 0800 111 757, or Youthline: 0800 37 66 33.

Project Contacts

Feel free to contact Edwin Chin (06 3569099 extn 7159, e.c.chin@massey.ac.nz) or the primary supervisor, Mei-Wah Williams (09 4140800 extn 41222, m.w.williams@massey.ac.nz) for further information or ask questions about the study, or if you have any issues or concerns.

This project has been reviewed and approved by the Massey University Human Ethics Committee: Northern, Application 10/035. If you have any concerns about the conduct of this research, please contact Dr Ralph Bathurst, Chair, Massey University Human Ethics Committee: Northern, telephone 09 414 0800 x 9570, email humanethicsnorth@massey.ac.nz.

Te Kunenga ki Pūrehuroa

School of Psychology – Te Kura Hinengaro Tangata
Private Bag 11222, Palmerston North 4442, New Zealand T 06 356 9099 extn 2040 F 06 350 5673 http://psychology.massey.

Appendix B – Consent Form

MASSEY UNIVERSITY
COLLEGE OF HUMANITIES
AND SOCIAL SCIENCES
TE KURA PŪKENGA TANGATA

New Zealand Survey of Students' Emotions

PARTICIPANT CONSENT FORM

I have read and understood the Information Sheet. I understand that my participation is voluntary and that I have the right to leave at any time.

I understand that I will be emailed by the researcher in January 2011, when I can choose whether to allow my English and Mathematics exam results for years 2009 and 2010 to be collected for this research. A summary of research findings will also be sent to my email.

My questions were answered to my satisfaction and I understand that I may ask further questions at any time. I agree to participate in this research under the conditions set in the Information Sheet.

Signature: _____ Date: _____

Full name: _____

Email: _____

Mobile: _____

Appendix C
– Study Questionnaire

- Reactions to Tests
- Reynolds Adolescent Depression Scale – II
- Positive and Negative Affect Schedule for Children
- Physiological Hyperarousal Scale for Children
- Screen for Child Anxiety Related Emotional Disorders – Revised
- Affect and Arousal Scale

New Zealand Survey of Student Emotions

A STUDY ON EXAM-RELATED WORRIES

There are five sections in this questionnaire.
Please read the instructions for each section carefully.

Do not spend too much time thinking about any particular question.
There are no right or wrong answers, just choose the option that best fits with you.

You have the right to not answer any particular question/s.

Your answers to this questionnaire are confidential.

Please respect the privacy of others around you by not looking at their answers.

Please circle your details below

age	16 / 17 / 18 / 19
year	12 / 13
gender	male / female
ethnicity	_____, _____ (please specify)

Please circle the NCEA subjects you are enrolled for <u>THIS YEAR.</u>

English	Level 1 / Level 2 / Level 3
Mathematics	Level 1 / Level 2 / Level 3 with Statistics / Level 3 with Calculus

Please circle the NCEA subjects you were enrolled for <u>LAST YEAR.</u>

English	Level 1 / Level 2 / Level 3
Mathematics	Level 1 / Level 2 / Level 3 with Statistics / Level 3 with Calculus

Section 1:
Think back to how you normally feel about tests and exams. For each of the items below, please circle the number that describes how typical or common it is for you to feel this way in an exam situation.

		Not Typical	Somewhat Typical	Quite Typical	Very Typical
1	I feel distressed and uneasy before tests.	1	2	3	4
2	The thought, "What happens if I fail this test?" goes through my mind during tests.	1	2	3	4
3	During tests, I find myself thinking of things unrelated to the material being tested.	1	2	3	4
4	I become aware of my body during tests (feeling itches, pain, sweat, nausea).	1	2	3	4
5	I freeze up when I think about an upcoming test.	1	2	3	4
6	I feel jittery before tests.	1	2	3	4
7	Irrelevant bits of information pop into my head during a test.	1	2	3	4
8	During a difficult test, I worry whether I will pass it.	1	2	3	4
9	While taking a test, I find myself thinking how much brighter the other people are.	1	2	3	4
10	I feel the need to go to the toilet more often than usual during a test.	1	2	3	4
11	My heart beats faster when the test begins.	1	2	3	4
12	My mind wanders during tests.	1	2	3	4
13	After a test, I say to myself, "It's over and I did as well as I could".	1	2	3	4
14	My stomach gets upset before tests.	1	2	3	4
15	While taking a test, I feel tense.	1	2	3	4
16	I find myself becoming anxious the day of a test.	1	2	3	4
17	While taking a test, I often don't pay attention to the questions.	1	2	3	4
18	I think about current events during a test.	1	2	3	4
19	I get a headache during an important test.	1	2	3	4
20	Before taking a test, I worry about failure.	1	2	3	4
21	While taking a test, I often think about how difficult it is.	1	2	3	4
22	I wish tests did not bother me so much.	1	2	3	4
23	I get a headache before a test.	1	2	3	4
24	I have fantasies a few times during a test.	1	2	3	4
25	I sometimes feel dizzy after a test.	1	2	3	4
26	I am anxious about tests.	1	2	3	4
27	Thoughts of doing poorly interfere with my concentration during tests.	1	2	3	4
28	While taking tests, I sometimes think about being somewhere else.	1	2	3	4
29	During tests, I find I am distracted by thoughts of upcoming events.	1	2	3	4
30	My hands often feel cold before and during a test.	1	2	3	4
31	My mouth feels dry during a test.	1	2	3	4
32	I daydream during tests.	1	2	3	4
33	I feel panicky during tests.	1	2	3	4
34	During tests, I think about how poorly I am doing.	1	2	3	4
35	Before tests, I feel troubled about what is going to happen.	1	2	3	4
36	The harder I work at taking a test, the more confused I get.	1	2	3	4
37	I sometimes find myself trembling before or during tests.	1	2	3	4
38	During tests I think about recent past events.	1	2	3	4
39	During tests, I wonder how the other people are doing.	1	2	3	4
40	I have an uneasy feeling before an important test.	1	2	3	4

Section 2:
Here are some sentences that people use to describe their feelings. Read each sentence and decide how often you feel this way, and tick the circle. Decide if you feel this way: almost never, hardly ever, sometimes, or most of the time.

		Almost Never	Hardly Ever	Sometimes	Most of the Time
1	I feel happy	o	o	o	o
2	I worry about school	o	o	o	o
3	I feel lonely	o	o	o	o
4	I feel my parents don't like me	o	o	o	o
5	I feel important	o	o	o	o
6	I feel like hiding from people	o	o	o	o
7	I feel sad	o	o	o	o
8	I feel like crying	o	o	o	o
9	I feel that no one cares about me	o	o	o	o
10	I feel like having fun with other students	o	o	o	o
11	I feel sick	o	o	o	o
12	I feel loved	o	o	o	o
13	I feel like running away	o	o	o	o
14	I feel like hurting myself	o	o	o	o
15	I feel that other students don't like me	o	o	o	o
16	I feel upset	o	o	o	o
17	I feel life is unfair	o	o	o	o
18	I feel tired	o	o	o	o
19	I feel I am bad	o	o	o	o
20	I feel I am no good	o	o	o	o
21	I feel sorry for myself	o	o	o	o
22	I feel mad about things	o	o	o	o
23	I feel like talking to other students	o	o	o	o
24	I have trouble sleeping	o	o	o	o
25	I feel like having fun	o	o	o	o
26	I feel worried	o	o	o	o
27	I get stomach aches	o	o	o	o
28	I feel bored	o	o	o	o
29	I like eating meals	o	o	o	o
30	I feel like nothing I do helps any more	o	o	o	o

There are no right or wrong answers. If you are unsure how to answer a particular item, feel free to ask the researcher.

Section 3:
Here is a list of different feelings and emotions that people have. Think about how much you have felt this way over the past few weeks, and put a number in the centre column next to the feeling.

(1) Not much or not at all. (2) A little. (3) Some. (4) Quite a bit. (5) A lot

Interested			Disgusted
Sad			Delighted
Frightened			Blue
Alert			Daring
Excited			Gloomy
Ashamed			Lively
Upset			Dry mouth
Happy			Sweaty hands / palms
Strong			Tingling (like pins and needles)
Nervous			Blushing
Guilty			Shaky
Energetic			Stomach ache
Scared			Cold flashes / chills
Calm			Dizzy
Miserable			Heart pounding
Jittery			Sweating when you are not hot
Cheerful			Can't catch your breath
Active			Feeling of choking
Proud			Hot flashes
Afraid			Numbness (like your foot's asleep)
Joyful			Pain in your chest
Lonely			Feeling like throwing up
Mad			Tight muscles
Fearless			Can't sit still

Section 4:
Below, you will find a number of statements, which refer to people's fears and anxiety. Please read each statement carefully and indicate how frequently you have experienced that symptom during the last 3 months: tick either 'never or almost never', 'sometimes', or 'often'.

		Never or almost never	Sometimes	Often
1	When I feel frightened, it is hard to breathe	o	o	o
2	I am afraid of heights	o	o	o
3	I get headaches or bellyaches when I am at school	o	o	o
4	I don't like to be with people I don't know	o	o	o
5	When I see blood, I get dizzy	o	o	o
6	I want things to be in a fixed order	o	o	o
7	I get scared when I sleep away from home	o	o	o
8	I worry about others not liking me	o	o	o
9	When I get frightened, I feel like passing out	o	o	o
10	I think that I will be contaminated with a serious disease	o	o	o
11	I am nervous	o	o	o
12	I have strange thoughts that frighten me	o	o	o
13	I follow my mother or father wherever they go	o	o	o
14	People tell me that I look nervous	o	o	o
15	I feel nervous with people I don't know well	o	o	o
16	I am afraid to visit the doctor	o	o	o
17	I don't like going to school	o	o	o
18	When I get frightened, I feel like I am going crazy	o	o	o
19	I worry about sleeping alone	o	o	o
20	I am afraid to visit the dentist	o	o	o
21	I worry about being as good as other kids	o	o	o
22	I am afraid of an animal that is not really dangerous	o	o	o
23	I get scared when there is thunder in the air	o	o	o
24	I do things more than twice in order to check whether I did it right	o	o	o
25	I have frightening dreams about a very aversive event I once experienced	o	o	o
26	I want things to be clean and tidy	o	o	o
27	When I get frightened, it feels like things are not real	o	o	o
28	I would feel scared if I had to fly in an aeroplane	o	o	o
29	I have nightmares about something bad happening to my parents	o	o	o
30	I worry about going to school	o	o	o

Remember: don't spend too much time thinking about a particular question. Simply pick the answer that best describes you.

	Continued from last section:	Never or almost never	Sometimes	Often
31	I perform rituals that help me to get less scared of my thoughts	o	o	o
32	When I feel frightened, my heart beats fast	o	o	o
33	I am scared when I get an injection	o	o	o
34	I am afraid of getting a serious disease	o	o	o
35	I feel weak and shaky	o	o	o
36	I have nightmares about something bad happening to me	o	o	o
37	I am so scared of a harmless animal that I do not dare to touch it	o	o	o
38	I worry about things working out for me	o	o	o
39	I doubt whether I really did something	o	o	o
40	When I get frightened, I sweat a lot	o	o	o
41	I am a worrier	o	o	o
42	I feel scared when I watch a medical operation on TV	o	o	o
43	I try not to think about a very aversive event I once experienced	o	o	o
44	Suddenly I get really frightened for no reason at all	o	o	o
45	I am afraid to be alone in the house	o	o	o
46	I get scared when I think back of a very aversive event I once experienced	o	o	o
47	It is hard for me to talk with people I don't know well	o	o	o
48	When I get frightened, I feel like I am choking	o	o	o
49	People tell me that I worry too much	o	o	o
50	I don't like to be away from my family	o	o	o
51	I am afraid of having anxiety (or panic) attacks	o	o	o
52	I worry that something bad might happen to my parents	o	o	o
53	I feel shy with people I don't know well	o	o	o
54	I have unwanted thoughts about hurting other people	o	o	o
55	I worry about what is going to happen in the future	o	o	o
56	When I get frightened, I feel like throwing up	o	o	o
57	I worry about how well I do things	o	o	o
58	I am scared to go to school	o	o	o
59	I worry about things that happened in the past	o	o	o
60	When I feel frightened, I get dizzy	o	o	o
61	I get scared in small, closed places	o	o	o
62	I have strange, scary thoughts that I prefer not to have	o	o	o
63	I am afraid of the dark	o	o	o
64	I have unbidden thoughts about a very aversive event I once experienced	o	o	o
65	I am afraid of an animal that most children do not fear	o	o	o
66	I don't like being in a hospital	o	o	o
67	I feel nervous when I have to do something while others are watching me (e.g., read aloud, speak, play a game, play a sport)	o	o	o
68	I feel nervous when I am going to parties, dances, or any place where there will people that I don't know well	o	o	o
69	I am shy	o	o	o

and finally, Section 5:
This section is about how you feel. For each sentence that you read, tick the answer that best tells how true that sentence is about how you usually feel. Remember, there are no right or wrong answers, just tick what you think describes you best.

		Never True	Sometimes True	Most Times True	Always True
1	When I 'm doing well at something, I really feel good	o	o	o	o
2	Other people upset me	o	o	o	o
3	Often I have trouble getting my breath	o	o	o	o
4	I get upset easily	o	o	o	o
5	My mouth gets dry	o	o	o	o
6	I have fun at school	o	o	o	o
7	My heart beats too fast	o	o	o	o
8	Little things bother me	o	o	o	o
9	I will try something new if I think it will be fun	o	o	o	o
10	My hands get shaky	o	o	o	o
11	When I get something I want, I feel excited	o	o	o	o
12	I over-react to things	o	o	o	o
13	I have trouble swallowing	o	o	o	o
14	I love going to new places	o	o	o	o
15	I get upset by little things	o	o	o	o
16	I feel shaky	o	o	o	o
17	I would love to win a contest	o	o	o	o
18	I don't like to wait for things	o	o	o	o
19	I like being with people	o	o	o	o
20	I have trouble breathing	o	o	o	o
21	When I see a chance for fun, I take it	o	o	o	o
22	I get upset	o	o	o	o
23	When good things happen to me, I feel full of energy	o	o	o	o
24	I have plenty of friends	o	o	o	o
25	I sometimes feel faint	o	o	o	o
26	I can't calm down once I am upset	o	o	o	o
27	Often I feel sick in my stomach	o	o	o	o

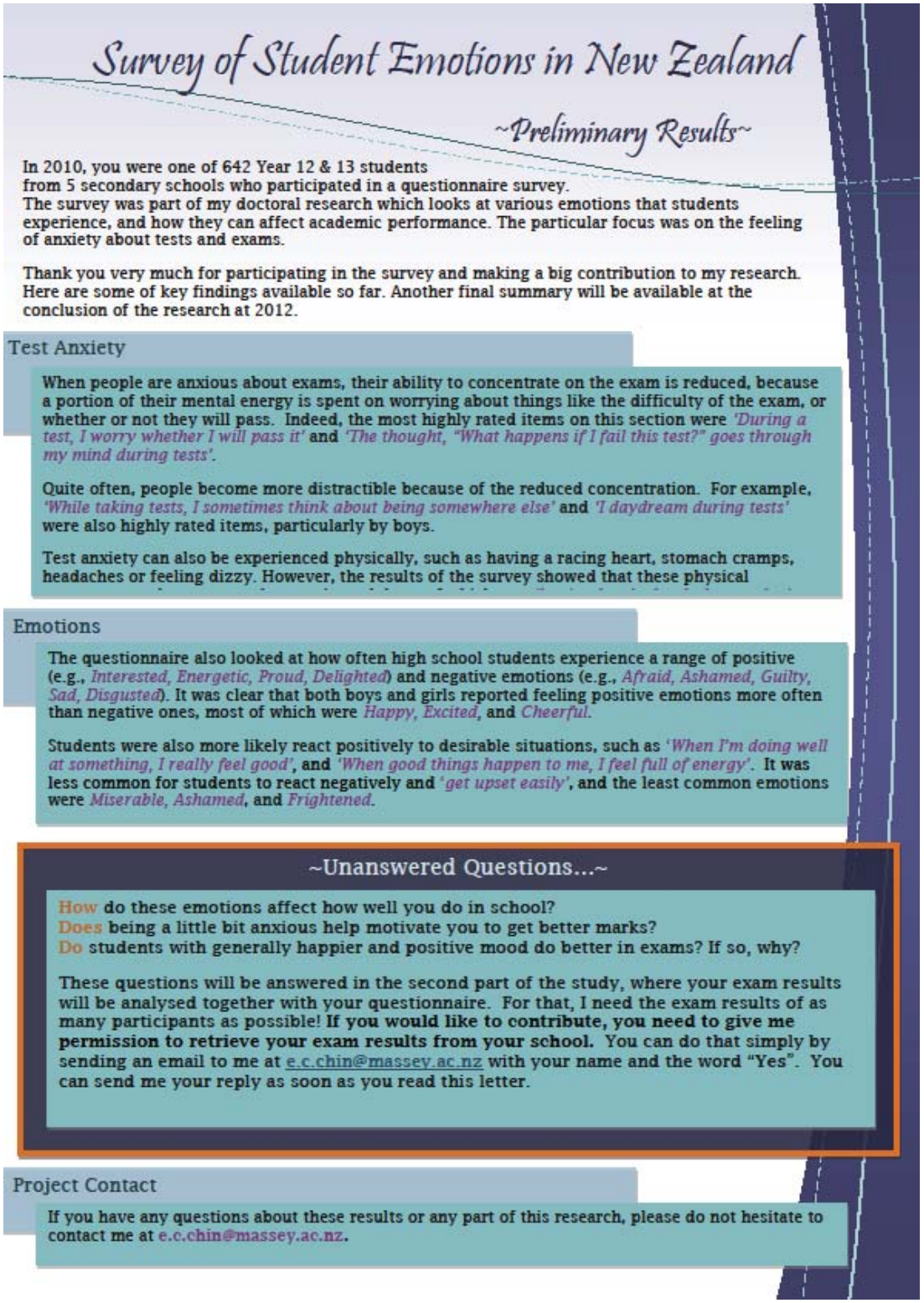
THAT'S IT!!

THANK YOU VERY MUCH FOR COMPLETING THE QUESTIONNAIRE

PLEASE RETURN THE QUESTIONNAIRE TO THE RESEARCHER.

IF YOU HAVE TAKEN THE QUESTIONNAIRE AWAY WITH YOU, YOU CAN POST IT BACK WITH THE FREEPOST ENVELOPE PROVIDED. IF YOU HAVE ANY QUESTIONS OR COMMENTS ABOUT THE STUDY, DO FEEL FREE EMAIL ME AT e.c.chin@massey.ac.nz.

Appendix D – Summary of Preliminary Results for Participants



Survey of Student Emotions in New Zealand

~Preliminary Results~

In 2010, you were one of 642 Year 12 & 13 students from 5 secondary schools who participated in a questionnaire survey. The survey was part of my doctoral research which looks at various emotions that students experience, and how they can affect academic performance. The particular focus was on the feeling of anxiety about tests and exams.

Thank you very much for participating in the survey and making a big contribution to my research. Here are some of key findings available so far. Another final summary will be available at the conclusion of the research at 2012.

Test Anxiety

When people are anxious about exams, their ability to concentrate on the exam is reduced, because a portion of their mental energy is spent on worrying about things like the difficulty of the exam, or whether or not they will pass. Indeed, the most highly rated items on this section were *'During a test, I worry whether I will pass it'* and *'The thought, "What happens if I fail this test?" goes through my mind during tests'*.

Quite often, people become more distractible because of the reduced concentration. For example, *'While taking tests, I sometimes think about being somewhere else'* and *'I daydream during tests'* were also highly rated items, particularly by boys.

Test anxiety can also be experienced physically, such as having a racing heart, stomach cramps, headaches or feeling dizzy. However, the results of the survey showed that these physical

Emotions

The questionnaire also looked at how often high school students experience a range of positive (e.g., *Interested, Energetic, Proud, Delighted*) and negative emotions (e.g., *Afraid, Ashamed, Guilty, Sad, Disgusted*). It was clear that both boys and girls reported feeling positive emotions more often than negative ones, most of which were *Happy, Excited,* and *Cheerful.*

Students were also more likely react positively to desirable situations, such as *'When I'm doing well at something, I really feel good'*, and *'When good things happen to me, I feel full of energy'*. It was less common for students to react negatively and *'get upset easily'*, and the least common emotions were *Miserable, Ashamed,* and *Frightened.*

~Unanswered Questions...~

How do these emotions affect how well you do in school?
Does being a little bit anxious help motivate you to get better marks?
Do students with generally happier and positive mood do better in exams? If so, why?

These questions will be answered in the second part of the study, where your exam results will be analysed together with your questionnaire. For that, I need the exam results of as many participants as possible! **If you would like to contribute, you need to give me permission to retrieve your exam results from your school.** You can do that simply by sending an email to me at e.c.chin@massey.ac.nz with your name and the word "Yes". You can send me your reply as soon as you read this letter.

Project Contact

If you have any questions about these results or any part of this research, please do not hesitate to contact me at e.c.chin@massey.ac.nz.

Appendix E – Email Summary of Final Results for Participants

Dear research participant

In 2010, you took part in a questionnaire survey that asked about how you experience positive and negative moods, as well as your thoughts and worries towards taking exams. The aim of the study was to explore the connections between general negativity and having exam-related worries. For example, do people who easily feel negative emotions (such as sad, jealous, ashamed, or guilty) also tend to feel worry about exams (such as the risk of failure and its consequences)? This is important because we know that exam-related worries can interfere with people's concentration during exams which reduces their performance.

The study found that the tendency to experience negative emotions influenced greater levels of exam-related worries. In turn, these exam-related worries were found to have impaired exam performance, but only to a small degree (i.e., not enough to drag you down a grade). Also, regardless of whether students are generally positive (excited and bubbly) or negative (a bit gloomy), that was not found to make a difference to exam performance.

Remember, it is normal for anyone to sometimes doubt their own academic abilities and worry about the personal, family, or social fallouts of failing exams. In fact, being a little anxious can help motivate people to study and do better. However, if those worries take over during an exam and they cause significant distress, a student counsellor may be able to help explore ways to reduce the anxiety.

If you have any questions about this study, feel free to contact me on Edwin.chin@gmail.com.

Many thanks for your participation in the survey and for making this project possible.

Edwin Chin
Doctoral candidate at the School of Psychology
Massey University

Appendix F – Residual Plots for Homoscedasticity

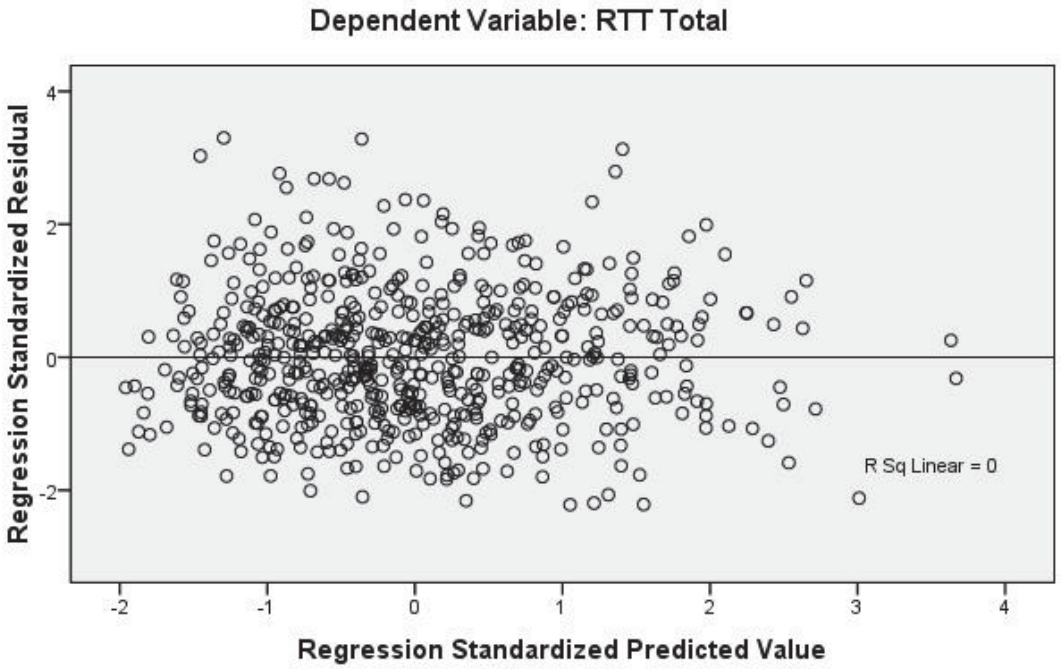


Figure F.1. Residual plot for cross-sectional sample.

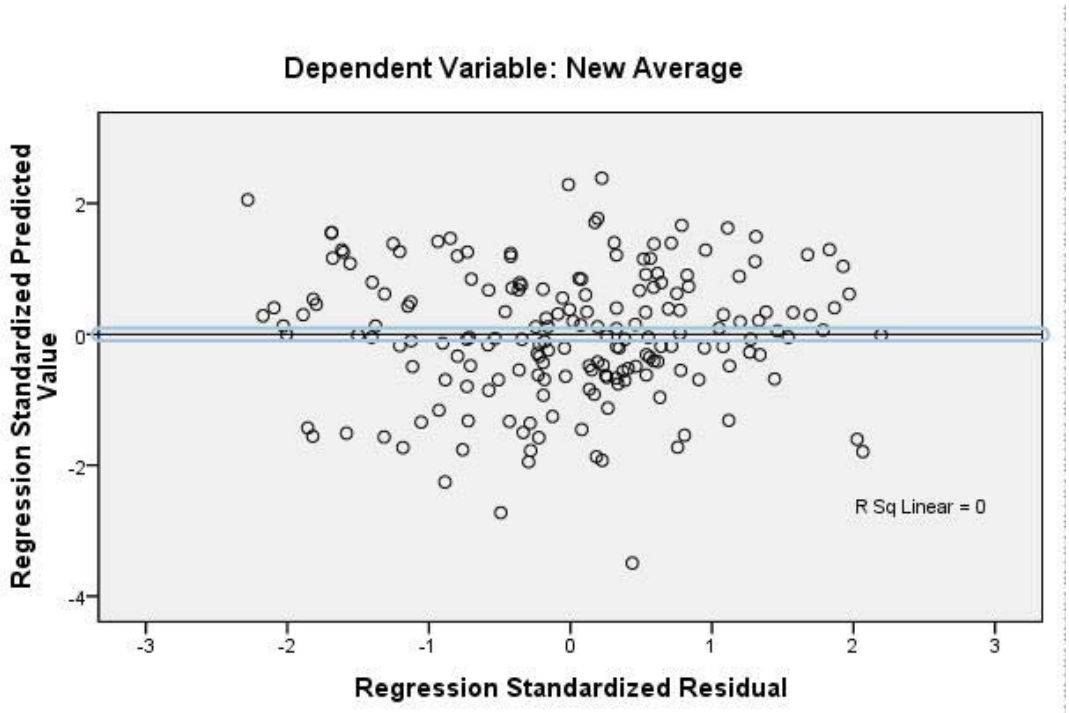


Figure F.2. Residual plot for the prospective sample.

Appendix G – Item Contents for the Trimmed Measurement Models

Table G

Content of Items Used to Represent to Latent Variables in Structural Modelling

Test Anxiety – (RTT)		
Worry	9:	While taking a test, I think how much brighter other people are.
	20:	Before taking a test, I worry about failure.
	21:	While taking a test, I often think about how difficult it is.
	34:	During tests, I think about how poorly I am doing.
Task-irrelevant thinking	7:	Irrelevant bits of information pop into my head during a test.
	12:	My mind wanders during tests.
	28:	While taking tests, I sometimes think about being somewhere else.
	32:	I daydream during tests.
Tension	6:	I feel jittery before tests.
	26:	I am anxious about tests.
	33:	I feel panicky during tests.
	40:	I have an uneasy feeling before an important test.
Bodily symptoms	11:	My heart beats faster when the test begins.
	23:	I get a headache before a test.
	31:	My mouth feels dry during a test.
	37:	I sometimes find myself trembling before or during tests.
Depression – (RADS-II)		
Anhedonia	1:	I feel happy.
	10:	I feel like having fun with other students.
	12:	I feel loved.
	25:	I feel like having fun.
Dysphoric mood	3:	I feel lonely.
	7:	I feel sad.
	8:	I feel like crying.
	16:	I feel upset.
Negative self-evaluation	9:	I feel that no one cares about me.
	15:	I feel that other students don't like me.
	20:	I feel I am no good.
	30:	I feel like nothing I do helps anymore.
Somatic complaints	11:	I feel sick.
	18:	I feel tired.
	24:	I have trouble sleeping.
	27:	I get stomach aches.
Anxiety – (SCARED-R)		
GAD	38:	I worry about things working out for me.
	41:	I am a worrier.
	49:	People tell me that I worry too much.
	57:	I worry about how well I do things.
Tripartite factors – (AFARS)		
Positive affect	1:	When I'm doing well at something, I really feel good.
	9:	I will try something new if I think it will be fun.
	19:	I like being with people.
	23:	When good things happen to me, I feel full of energy.
Negative affect	4:	I get upset easily.
	8:	Little things bother me.
	15:	I get upset by little things.
	22:	I get upset.
Physiological hyperarousal	7:	My heart beats too fast.
	13:	I have trouble swallowing.
	16:	I feel shaky.
	20:	I have trouble breathing.

Appendix H – Comparison of Tripartite Dimension Intercorrelations between Studies

Table H

Pearson's r Correlations of the Tripartite Factors for PH-PANAS and AFARS Across Comparable Studies

Between measures				Within measures		
PA-PA	NA-NA	PH-PH		PA-NA	PA-PH	NA-PH
Present study (N =617 Community adolescents)						
.43**	.57**	.57**	PH-PANAS-C	-.31**	-.06	.51**
			AFARS	.09*	-.12**	.55**
Jacques & Mash (2004) (N = 472 Community children and adolescents)						
.53***	.63***	.70***	PH-PANAS-C	-.31***	-.20***	.70***
			AFARS	-.15**	-.16***	.50***
Chorpita & Daleidin (2002) (N = 226 Clinical children and adolescents)						
.43**	.49**	N/A	PH-PANAS-C	-.29**	N/A	N/A
			AFARS	.00	-.20**	.55**
Stevanovic et al. (2013) (N = 449 Community adolescents)						
.34***	.36***	.66***	PH-PANAS-C	-.37***	-.17***	.51***
			AFARS	.11	-.18***	.36***

PA = Positive affect. NA = Negative affect. PH = Physiological Hyperarousal. PH-PANAS-C = Positive and Negative Affect Schedule for Children & Physiological Hyperarousal Scale for Children. AFARS = Affect and Arousal Scale. * $p < .05$. ** $p < .01$. *** $p < .001$