Distinguishing Low, Moderate, and High Test Anxiety

by

Taylor Crawford

A thesis

presented to the University of Waterloo

in fulfillment of the

thesis requirement for the degree of

Master of Arts

in

Psychology

Waterloo, Ontario, Canada, 2021

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

It is very common and understandable for post-secondary students to feel anxious leading up to a test. However, test anxiety can be a serious problem for some students. It is important to identify these individuals, given that test anxiety is associated with a variety of negative outcomes (Hembree, 1988; von der Embse et al., 2018). Although a plethora of self-report measures of test anxiety exist, arbitrary and inconsistent cut-off scores are currently used to make the distinction between those with varying levels of test anxiety (Huntley et al., 2019; Zeidner, 2014). It is evident that we require an empirically supported method of identifying those whose test anxiety is a problem, yet there is almost no research on what differentiates these individuals from those whose test anxiety is not problematic. The goal of the current study is to fill this gap in the literature by gaining a better understanding of the ways in which the experience of test anxiety may differ in those with low, moderate, and high severity levels. Undergraduate student participants were administered a semi-structured interview that inquired about their experiences before, during, and after testing situations. Participants also completed a commonly-used self-report measure of test anxiety (i.e., the Achievement Anxiety Test; Alpert & Haber, 1960). The interviewer separated participants into the severity groups according to an objective rating system based on the distress and interference their test anxiety was reported to cause. The information collected in the interview was then coded and qualitatively analyzed. The findings revealed many similarities in the thoughts, emotions, and bodily reactions individuals across groups experienced. However, some qualitative differences also emerged. Despite this, the questionnaire scores did not appear to distinguish participants in each of these groups from one another, demonstrating the measurement issues within the field. These results therefore have critical implications for the measurement, conceptualization, and treatment of test anxiety.

Acknowledgements

First and foremost, I want to thank my wonderful supervisor, Dr. Christine Purdon. I feel incredibly fortunate to learn from you every day. Completing my thesis during a global pandemic certainly brought its challenges, all of which I was able to navigate with your continual guidance and support. I would also like to thank my readers, Dr. Tara McAuley and Dr. Jonathan Oakman. Your insightful feedback provided me with many new and exciting considerations for my research both within and beyond this project. Many thanks are in order for the conscientious research assistants who spent hours transcribing and coding. Idiris Kabel, Aliya McNeil, Aliyah d'Gama Rose, and William Stuart: I truly appreciate all of the time and effort you devoted to this study. I am also grateful to have had the support of my inspiring cohort and lab, among all of the other members of the clinical program. This would not have been possible without the collaborative academic community of which I am honoured to be a part.

These acknowledgements would not be complete without recognizing my family. To my mom and dad: thank you for encouraging me to reach for my most ambitious goals and for always maintaining confidence that I will achieve them. To my partner, J: thank you for talking about my dreams as if they are your own, with nothing but excitement and trust and love. I would not be where I am today without all of you.

Table of Contents

List of Figures

List of Tables

PART I: LITERATURE REVIEW

Research on test anxiety is largely considered to have begun in the early 1950s, and our understanding of this construct has evolved significantly since then. Put simply, test anxiety encompasses feelings of fear or worry associated with testing situations (Bögels et al., 2010). Originally, research was focused on anxiety that is experienced before and during tests; however, it has expanded to include anxiety that arises afterwards as well (Mandler & Sarason, 1952; Bögels et al., 2010). According to Spielberger (1976), test anxiety, like anxiety in general, can be conceptualized in two ways. The Trait-State Theory of Anxiety differentiates between anxiety as an emotion that is experienced by all from time to time (i.e., state anxiety) and anxiety as a more stable characteristic of some individuals (i.e., trait anxiety) (Spielberger, 1972). It follows that almost everyone will experience state anxiety occasionally in testing situations; however, some students demonstrate a tendency to feel extremely anxious about testing. As such, Spielberger (1976) referred to test anxiety as a situation-specific trait. On the basis of this conceptualization, it has long been questioned whether test anxiety constitutes a disorder similar to other types of anxiety.

Test Anxiety as a Diagnosis

Many anxiety disorders are recognized in the most recent edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) (American Psychiatric Association, 2013). Although test anxiety is not currently acknowledged as a diagnosis, it was considered for inclusion in the DSM-IV (Bögels et al., 2010). Test anxiety was to be incorporated under Social Anxiety Disorder (SAD), with which it has been frequently theorized to be associated (McDonald, 2001; Zuriff, 1997). SAD is characterized by a fear of negative evaluation from others (American Psychiatric Association, 2013), and it has been suggested that this same fear is

at the core of test anxiety. Some studies have also demonstrated that children with SAD tend to endorse higher levels of test anxiety (Biedel & Turner, 1988; Biedel, Turner, & Trager, 1994). However, little research on comorbidity has been conducted with adults (Bögels et al., 2010).

Ultimately, test anxiety was not included in the DSM-IV due to issues regarding its conceptualization. The working definitions of test anxiety were capturing a large percentage of the population, rendering it impractical as a diagnosis (Bögels et al., 2010). Although there is little information publicly available regarding the information upon which their decision was based, it is reasonable to consider the possibility that the working group's estimate was exaggerated and the number of individuals who actually suffer from clinical test anxiety is lower. An overestimation may have occurred due to an overly vague or broad definition. Given that most individuals feel at least some anxiety about testing, a definition that lacks specificity has the potential to apply to a high percentage of individuals. It may also be the case that the working group's definition was inherently flawed as a result of test anxiety having been connected to the wrong disorder. Given that test anxiety was being considered for inclusion under SAD, the working definition may have largely focused on fear associated with the evaluative component of the test, rather than other factors (e.g., implications for the future). Although test anxiety has been most commonly associated with SAD in the literature, it has also been associated with Generalized Anxiety Disorder and thought of as a Specific Phobia (Bögels et al., 2010; LeBeau et al., 2010). Test anxiety may therefore be connected to various disorders for different individuals. It may also constitute a disorder in its own right, although this has not been investigated empirically.

Despite the current uncertainty with regard to how test anxiety fits within the context of other anxieties, it is important to note that test anxiety has the potential to be considered a

disorder if it manifests in a way that shares the common features of all of the anxiety disorders. This would be the case if the fear is clearly excessive, the feared stimulus is avoided or unpleasantly endured, and it causes significant distress and/or impairment in an important area of an individual's life (American Psychiatric Association, 2013; Bögels et al., 2010). Given that test anxiety has already been considered for inclusion in the DSM, it has likely already demonstrated that it can take this form for certain individuals, meaning it may reasonably be included in a future edition.

Components of Test Anxiety

Multiple elements of test anxiety have been identified within the literature, highlighting the various ways in which it can manifest. There are five components: cognitive, physiological, behavioural, affective, and social (von der Embse et al., 2018). Much of the research on these components focuses on their relation to indicators of academic success or lack thereof. The results of these studies have been amalgamated into two major meta-analyses, the results of which will be briefly summarized.

The two most well-established aspects of test anxiety are the cognitive and physiological components. These components were respectively referred to as worry and emotionality when they were first identified via factor analysis by Liebert and Morris (1967). The cognitive element is characterized by concern about the test, which involves worrying about one's performance and the consequences associated with it, among other stressors. The physiological element encompasses the typical bodily symptoms that are associated with feeling anxious, including but not limited to increased heart rate, difficulty breathing, stomachache, muscle tension, and sweating. Given that these are the longest-standing components of test anxiety, they have accumulated a great deal of research. Although both are important manifestations, worry has

been consistently shown to be more strongly associated with negative outcomes, such as decreased test performance (Hembree, 1988; von der Embse et al., 2018).

The behavioural component of test anxiety has also been steadily researched. When experiencing anxiety of any kind, individuals are motivated to reduce that unpleasant emotion. Anxiety can be reduced in two main ways. First, one may tackle the subject of the anxiety (e.g., studying for the test), which reduces anxiety by addressing its source. Second, one might engage in task-irrelevant behaviour (e.g., watching television), which reduces anxiety by distracting from or avoiding its source for a short time. The latter option constitutes the behavioural component. Interestingly, Von der Embse et al.'s (2018) recent meta-analysis showed that this component was not significantly associated with performance deficits. This suggests that distraction from studying in and of itself is not necessarily the culprit of declining grades, as one might reasonably conclude without this finding.

The affective component of test anxiety refers to the feeling of anxiety itself. As such, all research on test anxiety inherently involves this component. Despite the fact that anxiety is by far the most-researched emotion connected to test-taking, Pekrun et al. (2004) remind us that a wide variety of feelings are associated with tests and that these can arise concurrently with anxiety. The affective component therefore encompasses all of the emotions that are experienced while feeling anxious about a test. Although the emotions most closely associated with test anxiety are negatively-valanced (e.g., shame, hopelessness), Pekrun et al. (2004) demonstrate that positively-valanced emotions (e.g., joy, pride) are also related to anxiety to a lesser extent. Interestingly, hopelessness is more closely associated with low test performance than test anxiety itself. This research therefore underscores the importance of conceptualizing the affective component of test anxiety in a broad sense.

The most recently recognized aspect of test anxiety is the social component. This is unsurprising given the aforementioned theorized relation between test anxiety and SAD. Thus far, this component has been most widely acknowledged in children; however, it may also be applicable to adults (Lowe et al., 2008; Segool et al., 2014). The social component is an extension of the cognitive component because it manifests itself in terms of thoughts regarding comparison as well as worries about negative evaluation and subsequent degradation from others (e.g., family, friends, teachers). Researchers have considered this aspect distinct enough to be acknowledged in and of itself. Like the cognitive component, the social component has been shown to be related to decreased test performance, though not quite as strongly (von der Embse et al., 2018). With all that being said, more research on the social worries of test anxiety is required given that its introduction as a component in its own right is still relatively new.

Measurement of Test Anxiety

As theories and research regarding the different components of test anxiety emerged, so too did measures of this construct. By far, the most common way to measure test anxiety is through self-report questionnaires, and a plethora of these have been developed. The first well-recognized measure of test anxiety was created by Mandler and Sarason in 1952. A number of measures have since been derived from this original instrument, titled the Test Anxiety Questionnaire (TAQ). These measures will be outlined first, followed by several other measures that have been independently constructed. It is important to note that although the TAQ is largely considered to be the first measure of test anxiety, it was in fact preceded by an unnamed questionnaire that was published more than a decade earlier (Brown, 1938). This signified the true beginning of research on test anxiety, although it was not referred to as such at the time. Given that this unnamed measure is quite far-removed from the other measures and does not

seem to have had an influence on their construction, it will not be outlined in full. However, it is worth acknowledgement from a historical perspective. It is also important to note that although only the measures intended for use with adults will be outlined, several self-report measures of test anxiety intended for use with children and/or adolescents also exist (e.g., the Test Anxiety Inventory for Children and Adolescents) (Lowe et al., 2008).

Related Self-Report Measures

The TAQ was originally developed to facilitate Mandler and Sarason's (1952) research on drive states. This was the first study to investigate the relationship between anxiety and test performance. The 67 original questions on the TAQ ask about individuals' attitudes toward testing and their experiences with individual and group intelligence tests as well as university examinations. Although the components of test anxiety had not yet been identified, the questions would fall within today's cognitive (e.g., worries), physiological (e.g., sweating), and emotional (e.g., uneasiness) domains. Participants answered each question on a graphic scale, and their responses were scored and summed to deliver a total score for test anxiety. The TAQ was commonly used by researchers for decades, and although it is not frequently selected today, it has been survived by a variety of other measures that can directly or indirectly be traced back to the TAQ.

The first of these was Sarason's Test Anxiety Scale (TAS) (1958). The original version of the TAS was composed of a subset of 21 items taken directly from the TAQ. Sarason changed the format of the questions to be true or false for ease of use. Like the TAQ, the TAS generates a total test anxiety score. Upon further analysis of the TAS, Sarason (1972) created a subsequent version with 37 items, which included some novel content. This version of the TAS remains quite popular today.

While Sarason worked on the revised version of the TAS, Liebert and Morris (1967) also took interest in the TAQ. From their examination, two factors were identified: worry (i.e., the cognitive component) and emotionality (i.e., the physiological component). To explicate which of these two factors were most closely related to test performance, Liebert and Morris selected 10 items from the TAQ (5 of which they regarded to fall under worry and 5 of which they regarded to fall under emotionality), to create what is referred to as the Worry-Emotionality Questionnaire. Participants respond to items on a 5-point Likert scale, and their answers are summed to create a score for each of the factors. Although the Worry-Emotionality Questionnaire was not commonly-used beyond this study, the theory upon which it was based inspired other measures. This theory has clearly withstood the test of time considering that the cognitive and physiological components of test anxiety are still the most widely researched.

Osterhouse's Inventory of Test Anxiety (1972) and Spielberger et al.'s Test Anxiety Inventory (TAI) (1978) were two of the measures that adopted the structure of the Worry-Emotionality Questionnaire. Osterhouse's 16-item inventory, which was composed of questions from multiple other measures, did not garner significant attention, unlike that of Spielberger et al. The TAI was constructed shortly after the first author's development of the Trait-State Theory of Anxiety (1972), which was outlined previously. Consistent with this theory, the TAI was intended to measure test anxiety as a situation-specific trait. The authors borrowed items from the TAS and augmented the question set with their own items. Respondents answer the 20 questions on Likert scales, and the inventory yields worry, emotionality, and total scores. The TAI became and has remained one of the most frequently-used measures of test anxiety, and it has been translated into multiple languages (e.g., German, Japanese).

As updated theories regarding test anxiety and its components were published, Sarason seemed to have noticed the value in creating another measure of test anxiety that included multiple scales. This was in contrast to both versions of the TAS which produce only a total score. This new questionnaire, titled Reactions to Tests (RTT) (1984), includes items from the TAS in addition to many novel items, which individuals respond to on a 4-point scale. The RTT has four 10-item scales: worry, test-irrelevant thinking, bodily symptoms, and tension. Sarason suggested that each of these represent a distinct component of test anxiety. The first and second scales refer to what we now call the cognitive component. They are not currently conceptualized separately, although test-irrelevant thinking alludes to the behavioural component in a way (i.e., task-irrelevant behaviours or distraction). Sarason stated that the third and fourth scales were constituents of emotionality, with bodily symptoms referring to specific physiological reactions (e.g., stomachache) and tension referring to more general feelings (e.g., panicky). The latter seems to have been the first recognition of a distinct affective component of test anxiety (Pekrun et al., 2004). Decades later, this structure was embraced by other researchers.

Benson & El-Zahhar (1994) adopted the RTT's four factors for their questionnaire, the Revised Test Anxiety Scale (RTAS). In addition to taking items from the RTT, the RTAS also included items from the TAI. The final version of the RTAS is comprised of 20 items, some of which were created by the authors. Pekrun et al. (2004) was also heavily influenced by the RTT when constructing their test anxiety scale, the Test Emotions Questionnaire (TEQ). As mentioned, these authors played an integral role in establishing the affective component by highlighting that anxiety is accompanied by other emotions in testing situations. The TEQ is intended to assess a variety of feelings associated with testing. It was first written in German and later translated to English. The final version has six scales: joy, relief, anger, anxiety, shame, and

hopelessness. Each scale also includes affective, cognitive, physiological, and motivational subscales. The authors mapped the first three subscales of their anxiety scale onto the RTT's tension, worry, and bodily symptoms scales, respectively. The RTT's items were also used for these subscales. For the motivational subscale of the TEQ, new items that related to escaping testing situations were constructed. In this way, the motivational subscale of anxiety on the TAQ overlaps with the behavioural component of test anxiety. To the author's knowledge, the TEQ is the most recent measure that can ultimately be traced back to the TAQ.

Independent Self-Report Measures

Although many measures of test anxiety stemmed from the TAQ, several were also constructed independently. The Achievement Anxiety Test (AAT) was the first of these. Alpert and Haber (1960) took a different approach than that of other researchers in that the main focus of their measure is on the impact of test anxiety. The AAT has a 10-item debilitating scale (AAT-D) and a 9-item facilitating scale, which respectively estimate the extent to which test anxiety hinders or assists with performance. The AAT-D quickly became a popular choice among test anxiety researchers, and it is still used today.

Years later, the Suinn Test Anxiety Behaviour Scale (STABS) was published (Suinn, 1969). By this time, research on test anxiety had taken off and treatments were being studied. The STABS was intended to evaluate the effectiveness of one such intervention, namely behaviour therapy, which is alluded to in the name of the measure. The STABS is composed of 50 questions which enquire about the extent to which different testing-related situations incite anxiety. It is still occasionally used today.

It was decades before another noteworthy measure of test anxiety independent from the TAQ was created. In 2001, Cassady and Johnson published the Cognitive Test Anxiety Scale

(CTAS). As its' name suggests, the CTAS is concerned with only the cognitive component of test anxiety. Although its 27 items were developed independently, the CTAS measures task-irrelevant thinking, similar to one of the scales on the RTT. It has also been compared to the TAI. In 2018, Thomas, Cassady, and Holmes Finch published the second version of the CTAS. To the author's knowledge, these are the only test anxiety researchers to have attempted to establish cut-off scores for their measure. The CTAS-2 is thus the sole measure of test anxiety with a set of preliminary cut scores intended to identify those with low, moderate, and high test anxiety.

In 2007, Driscoll constructed the Westside Test Anxiety Scale (WTAS) as a means by which to identify individuals in need of intervention for their test anxiety. Similar to Alpert and Haber (1960), the WTAS includes six items regarding the impairment caused by test anxiety. Its other four items focus on worry (i.e., the cognitive component), which has been shown to be associated with performance deficits. The brevity of the WTAS renders it a popular choice among researchers.

Problems with the Measurement of Test Anxiety

Although several of the aforementioned questionnaires are still widely used today, a number of issues with test anxiety measurement have been identified. The overarching problem is that these measures do not seem to meaningfully distinguish those whose test anxiety is a problem from those whose test anxiety is not. Questionnaires that measure recognized anxiety disorders (i.e., those included in the DSM-5) have established cut-off scores to differentiate those with clinical vs. sub-clinical anxiety. The vast majority of test anxiety measures do not have cut-offs that are similar in nature. Although the CTAS-2 has preliminary cut scores, these have yet to be validated or used, to the author's knowledge, outside of Thomas,

Cassady, and Holmes Finch's (2008) investigation. Despite all of this, researchers often impose cut-off scores onto test anxiety measures in order to select those with "high" test anxiety for their studies. Unfortunately, the scores are selected arbitrarily and inconsistently across a variety of measures. Some researchers use a raw questionnaire score as a cut-off, whereas others use a percentile (Huntley et al., 2019). The specific raw score or percentile that is selected can vary greatly. As such, any two given studies can define "high" test anxiety differently (e.g., >30 raw score on the AAT vs. >80th percentile on the TAS). Although it seems likely that there is a meaningful distinction to be made between those whose test anxiety is a problem (i.e., clinical) and those whose test anxiety is not (i.e., subclinical), this differentiation is not currently being made in research.

The ways in which these measures continue to be utilized by researchers is unfortunately not where the issue ends. It is entirely possible that it is invalid to apply any cut-off score to many of the existing test anxiety measures. As mentioned, these measures are largely concerned with the attitudes and symptoms individuals experience in the context of testing. However, it is not necessarily symptoms that differentiate subclinical vs. clinical anxiety, but rather the excessiveness of those symptoms as well as the distress and/or interference they cause (American Psychiatric Association, 2013). As such, the symptoms that are queried in the test anxiety measures may very well be similar for those with low, moderate, and high test anxiety. For example, one of the questions included on the TAS (1972, pp. 384) reads, "while taking an important examination I perspire a great deal." Writing tests is expected to cause some anxiety or stress in adults, given that they often have academic, vocational, or other important implications. Sweating is a typical response to even low levels of anxiety. As such, those who have low test anxiety overall may still sweat while writing a test. Therefore, individuals with varying severity

levels may answer questions like this similarly, which may essentially lump everyone with test anxiety together in the end. This is not entirely surprising when the history of these measures is considered.

Recall that many of the questionnaires that are still in use today can be linked back to one of the very first measures of test anxiety, the TAQ (Mandler & Sarason, 1952). The study in which the TAQ was published is largely thought to signify the beginning of test anxiety research. This means that the questions on the TAQ, which have also been carried on to newer measures, were constructed before much (if anything) was truly known about test anxiety. This calls into question the backbone of many of the existing test anxiety measures. Although the questionnaires that did not originate from the TAQ do not include its potentially outdated items, several of these measures remain similar in nature in the sense that at least some of their questions also revolve around attitudes and symptoms associated with test-taking. Even the measures that claim to focus on impairment (e.g., AAT, WTAS) include questions of this nature. In sum, there is little variety in the types of questions that current test anxiety instruments ask of respondents outside of these two major themes, and these questions may not be reflective of actual differences between those with varying test anxiety severity levels.

In addition to the individualized criticisms outlined above, more general concerns about self-report questionnaires are also relevant to the evaluation of test anxiety measures. Although self-report instruments are convenient, they are also subject to biases and intentional or unintentional suppression or exaggeration on the part of the respondent (Zeidner & Matthews, 2003). Limited response options (e.g., true or false) can generate misleading results and although Likert scales attempt to capture nuance, it can be very challenging for respondents to consider where they might fall. It is likely difficult for individuals to evaluate how their anxiety compares

to that of others. One student who is objectively much more anxious about tests may believe their experience is similar to others' experiences and thus rate themselves lower. On the other hand, a student who is objectively much less anxious about tests, may believe their experience is typical or even more extreme than others' experiences and rate themselves higher as a result. Alternatively, students may simply select a neutral response option by default.

Despite these additional considerations, it makes sense that test anxiety is most commonly measured using self-report. Although some researchers have employed other methods for its assessment, they are certainly more time-consuming and costly. For example, Reiss et al. (2017), used DSM-IV criteria for either SAD or SP to identify individuals with test anxiety via the Structured Clinical Interview for DSM-IV. This interview likely would have required the researchers to dedicate 1-2 hours individually with each participant. In contrast, self-report measures typically take less than 10 minutes, can be completed concurrently by many participants, and do not necessarily require the researcher to be present.

Overall, several issues currently exist with test anxiety measurement. Although these problems have been acknowledged for quite some time, they have all remained largely unaddressed (Zeidner, 2014). It is important to consider the research and clinical implications of such longstanding issues.

Implications of the Problems with Test Anxiety Measurement

Treatment of Test Anxiety

The literature surrounding test anxiety treatments has grown immensely over the last few decades. The three most commonly used treatments for test anxiety are cognitive-behavioural therapy, behaviour therapy, and study skills training (Huntley et al., 2019). The first two of these are psychological interventions, which address the anxiety directly. The final intervention aims

to equip students with the skills required to improve their studying and test-taking abilities in an effort to indirectly reduce anxiety. Huntley et al.'s (2019) meta-analysis on the efficacy of these interventions demonstrated that each one significantly reduces test anxiety; however, the authors cautioned that their results must be interpreted carefully due to the measurement issues within the field. Nearly 10 different questionnaires were used to assess test anxiety across the studies included. Of the 44 studies, 30 used raw scores or percentiles as cut-offs to identify those with increased or "high" test anxiety. The chosen cut-offs differed considerably from one another (e.g., 50^{th} to 85^{th} percentile) across studies. The remaining 14 studies did not use cut-offs to screen participants, meaning they likely included individuals of varying severity levels. The astounding variability in measurement makes it challenging to interpret the results of these studies and make meaningful conclusions because it is unclear whom these interventions were tested on. Thus, the generalizability of these findings is called into question.

Accommodations for Test Anxiety

Despite the disadvantage to which those with consistently and remarkably impairing test anxiety are subjected, numerous post-secondary institutions and testing organizations do not recognize test anxiety as being sufficient for accommodations (e.g., University of Western Ontario, Educational Testing Service). In order to receive accommodations, an individual oftentimes must have a documented diagnosis. Such requirements were likely established to ensure that only those students who truly need accommodations are able to receive them. However, as mentioned, test anxiety is not classified as an anxiety disorder in the DSM-5. Although this is largely a conceptualization issue, measurement certainly acts as a barrier to achieving this clarification and implementing accommodations, as the measures that are currently available do not meaningfully distinguish between those whose test anxiety is an issue

and those whose test anxiety is not. A significant amount of research is still required to fully understand this phenomenon and address issues within the field.

PART II: EMPIRICAL STUDY

Introduction

Feeling anxious about a test from time to time is a very common and understandable experience for post-secondary students. Although some anxiety can be harmless, and in some cases even helpful, too much can be a hinderance. Increased test anxiety has been consistently shown to be associated with impaired performance on course examinations, standardized tests, and achievement in school more broadly (Hembree, 1988; von der Embse et al., 2018). Given that tests are unavoidable in our society and test scores are often used to make critically important life decisions, those who frequently suffer from debilitating test anxiety are at a significant disadvantage. It is of great importance to be able to identify such individuals to facilitate the implementation of treatment; however, test anxiety is not recognized as a disorder with established criteria and existing measures of test anxiety do not necessarily distinguish between those whose test anxiety is a problem (i.e., high severity) and those whose test anxiety is not (i.e., low or moderate severity). This reflects a large gap in our understanding of this phenomenon that has endured in spite of decades worth of research.

Construct of Test Anxiety

Test anxiety encompasses feelings of fear or worry associated with testing situations (Bögels et al., 2010). It can occur before, during, and/or after the administration of the test. Within the context of his Trait-State Theory of Anxiety, Spielberger suggested that test anxiety is a situation-specific trait (Spielberger, 1972; Spielberger, 1976). Although almost everyone may experience anxiety occasionally in testing situations, some individuals demonstrate a tendency to feel extremely anxious under these circumstances (Spielberger, 1976). As such, researchers and clinicians have continually explored whether or not test anxiety may constitute a disorder.

Although test anxiety was being considered for inclusion under Social Anxiety Disorder (SAD) in the DSM-5, the decision to exclude it was ultimately made (American Psychiatric Association, 2013; Bögels et al., 2010). This was due to a vague working definition of test anxiety that applied to a large number of individuals as well as uncertainty in terms of whether test anxiety truly fits within SAD (Bögels et al., 2010). Our conceptualization of test anxiety has yet to advance very far beyond these considerations.

Fortunately, we do seem to have a fair understanding of the ways in which test anxiety can manifest. Five components have been identified within the literature, namely cognitive, physiological, behavioural, affective, and social. Liebert and Morris (1967) introduced the cognitive and physiological components first. The former is characterized by concern about the test (e.g., grades, academic consequences), whereas the latter encompasses a variety of bodily reactions to anxiety (e.g., increased heart rate, difficulty breathing). The behavioural element refers to the indirect actions that an individual takes to manage their anxiety (e.g., avoidance, distraction) (von der Embse et al., 2018). The affective component is the feeling of anxiety itself, in addition to other emotions that are associated with test-taking (e.g., shame, hopelessness, joy) (Pekrun et al., 2004). Finally, the social component involves worrying about negative evaluation and degradation from others as a result of test performance (von der Embse et al., 2018). The experience that an individual has with test anxiety may involve any combination of these components, and their intensity may shift both within and between any given testing situation. As such, test anxiety may be exhibited in a wide variety of ways.

Measurement of Test Anxiety

A plethora of self-report measures of test anxiety currently exist. The first well-recognized measure was constructed by Mandler and Sarason (1952). Their study is generally

considered to signify the beginning of test anxiety research. The Test Anxiety Questionnaire (TAQ) is composed of items written by the researchers regarding individuals' attitudes toward testing and the symptoms experienced during tests. Despite that this original questionnaire was based on little (if any) empirical research, the majority of the questionnaires that have since been established were derived from the TAQ. In 1958, Sarason changed the format of a subset of the TAQ's questions to create the Test Anxiety Scale (TAS). The TAS was later expanded upon with the addition of novel questions, and this revised version is still commonly-used (1972). Between these versions of the TAS, Liebert and Morris (1967) analyzed the TAQ and discovered two factors that would come to be known as the cognitive and physiological components of test anxiety (i.e., worry and emotionality). The Worry-Emotionality Questionnaire is composed of 10 questions that were adapted from the TAQ, which contribute equally to a cognitive and an affective score. Osterhouse's Inventory of Test Anxiety (1972) and Spielberger's Test Anxiety Inventory (TAI) (1978) adopted this structure for their questionnaires. The latter remains one of the most frequently-used measures of test anxiety. In 1984, Sarason published another test anxiety questionnaire, titled Reactions to Tests (RTT), the structure of which reflected updated theories. The RTT has four scales that directly or at least loosely map onto some of the components of test anxiety that are recognized today: worry, test-irrelevant thinking, bodily symptoms, and tension. This four-factor model was used by Benson & El-Zahhar for their Revised Test Anxiety Scale (RTAS) (1994), which also borrowed items from several other measures. The RTT also heavily influenced the development of Pekrun et al.'s anxiety scale on the Test Emotions Questionnaire (TEQ) (2004). Test-related joy, relief, anger, shame, and hopelessness are measured by the other five scales. To the author's knowledge, the TEQ is the most recently published measure that can be traced back to the TAQ.

A number of other test anxiety measures have been constructed independently over the last several decades. First, the Achievement Anxiety Test (AAT) was created by Alpert and Haber in 1960. Instead of focusing solely on the attitudes and symptoms associated with testing, like the majority of the measures that were derived from the TAQ, the AAT focuses on the impact of test anxiety. It is composed of a facilitating and a debilitating scale. As its name suggests, Cassady and Johnson's Cognitive Test Anxiety Scale (CTAS) measures only the cognitive component of test anxiety, instead of attempting to capture multiple aspects of the experience (2001). The second version of the CTAS (2018) is the only measure of test anxiety that has a set of preliminary cut-off scores to identify those with low, moderate, and high test anxiety. However, these cut-offs have yet to be validated or used, to the author's knowledge, outside of Thomas, Cassady, and Holmes Finch's (2008) investigation. Driscoll's (2007) Westside Test Anxiety Scale also focuses on the cognitive component, as well as the impairment caused by test anxiety. It is a frequently selected measure due to the ease with which it can be administered. It is important to note that other test anxiety measures exist outside of this compilation; however, their use is more restricted (e.g., Suinn Test Anxiety Behaviour Scale) or aimed at children and adolescents (e.g., the Test Anxiety Inventory for Children and Adolescents) (Suinn, 1969; Lowe, 2008).

Although many of these questionnaires are widely used today, several issues have been identified with test anxiety measurement. The overarching problem is that these measures do not seem to meaningfully distinguish those whose test anxiety is a problem from those whose test anxiety is not. As mentioned, these measures are largely concerned with individuals' attitudes toward testing and the symptoms they experience during tests. The problem with this is that these attitudes and symptoms are likely similar for those with varying levels of test anxiety, meaning

many individuals may answer questions in the same way. This would essentially lump everyone with test anxiety together instead of differentiating based on severity. Although it may therefore be inappropriate to use cut-off scores on the questionnaires to make this distinction and cut-offs have not been empirically validated for the vast majority of these measures, researchers often impose arbitrary-selected cut-offs to identify those with "high" test anxiety (Huntley et al., 2019; Zeidner, 2014). Consequently, any two given studies may define high test anxiety differently (e.g., >30 raw score on the AAT vs. >80th percentile on the TAS). There are several significant implications of these measurement issues. First, it is challenging to interpret the results of studies that investigate test anxiety, as it is unclear to whom the results apply when we lack a common understanding of high test anxiety. For example, although several treatments (e.g., cognitive-behavioural therapy, behaviour therapy, study skills training) have been found to effectively reduce test anxiety, the generalizability of these findings must be considered with caution (Huntley et al., 2019). Second, numerous post-secondary institutions and other testing organizations across Canada do not recognize test anxiety as being sufficient for academic accommodations. This is likely more so due to issues with the conceptualization of test anxiety (i.e., uncertainty with regard to whether or not test anxiety constitutes a diagnosis), but measurement undoubtedly acts as a barrier to the implementation of such accommodations.

Current Study

The issues that have been identified with test anxiety measurement are only continuing to grow. It is evident that we require an empirically supported method of identifying those with high test anxiety. The reason such a measure has not been created is that there is almost no research on what distinguishes these individuals. The current study primarily aimed to fill this gap in the literature by exploring how the experience of test anxiety differs in those who were

objectively determined to have low, moderate, and high levels of test anxiety. The research questions are as follows:

1. How do the thoughts associated with test anxiety differ among the low, moderate, and high test anxiety groups?
2. How do the emotions associated with test anxiety differ among the low, moderate, and high test anxiety groups?
3. How do the bodily reactions associated with test anxiety differ among the low, moderate, and high test anxiety groups?

To the author's knowledge, this study was the first to distinguish participants who experience low, moderate, and high test anxiety according to objective interviewer ratings. This affords the opportunity to compare test anxiety questionnaire scores across groups in an effort to evaluate the questionnaire's effectiveness in distinguishing severity. A secondary aim of the study was therefore to evaluate one of the current measures of test anxiety. The final research question is as follows:

4. Does the Debilitating Scale on the Achievement Anxiety Test (AAT) differ among the low, moderate, and high test anxiety groups?

Method

Participants

Thirty-eight undergraduate students from the University of Waterloo participated in this study. Two participants were excluded from the final sample as they did not complete the interview in full. This was due to technical difficulties for one student and restricted time for the other. Thirty-six participants therefore comprised the final sample. The number of participants in the final sample fell within the range (30-40) that the author was originally aiming for based on Johanson and Brooks' (2010) recommendation.

Students were recruited through the university's SONA system. An initial screening for anxiety occurred using the Depression Anxiety and Stress Scale – 21 (Lovibond & Lovibond, 1995). This was administered to SONA's participant pool as a component of the Mass Testing and Prescreen Questionnaire. In order to be eligible, students needed to obtain a score of 18 or higher on the stress subscale only. This cut-off represents a score that is 0.5 standard deviations below an established mean for individuals with Generalized Anxiety Disorder (Brown et al., 1997). It was selected as the cut-off because students with lower scores would have been less likely to experience any test anxiety at all.

To participate, students also needed to have experienced test anxiety within the last 12 months at university. This information was included on the recruitment materials that individuals who made it through the screen were able to view on SONA. The description of test anxiety on the recruitment materials was left intentionally vague in an effort to encourage those with varying levels of test anxiety severity to sign-up for the study. Following completion of the study, the author separated participants into groups using an objective rating system. This rating system will be outlined further in the method section.

Measures

Test Anxiety Interview

A semi-structured interview, titled the Test Anxiety Interview, was constructed in order to gain a better understanding of individuals' experiences with test anxiety and thereby answer the primary research questions. The structure of the interview was based on Chiang and Purdon's (in prep) Phenomenological Interview of the Obsessive-Compulsive Experience. The Test Anxiety Interview is broken down into six sections. The first five sections include questions that are focused around one particular instance of test anxiety. The first section is intended to orient the interviewee and the interviewer to the selected instance of test anxiety by obtaining background information. Information about the test (e.g., length, format) and the course for which the test was completed (e.g., mandatory or elective course, grading breakdown) is collected. In the second section, the interviewee is asked to describe their experience of anxiety before, during, and after the test. Aligning with the first three research questions, this section includes inquiries about individuals' thoughts, emotions, and bodily reactions across these time periods. The third and fourth sections of the interview are based on Lazarus and Folkman's Transactional Model of Stress and Coping (1984). According to this model, stress or anxiety manifests when an individual appraises the demands of their environment (e.g., importance of a test) to exceed their ability to cope (e.g., retain information). It is important to note that the data collected within these sections will not be included within this thesis. The fifth section focuses on the distress and interference test anxiety causes across various domains (e.g., academic, social, health and well-being). The sixth and final section includes broader questions that ask about test anxiety in general (e.g., frequency, intensity, onset, course). The full interview can be found in the Appendix. This interview was completed individually with the undergraduate

student participants. The information collected in various sections of the interview was utilized differently. For this thesis, the data from the second section on thoughts, emotions, and bodily reactions was coded and qualitatively analyzed, and the information from the fifth and sixth sections was used to assign each participant a test anxiety severity score which separated individuals into groups.

Test Anxiety Questionnaire

Alpert and Haber's (1960) Achievement Anxiety Test (AAT) was the self-report measure selected for administration in order to answer the secondary research question. This questionnaire was chosen because it is one of the few test anxiety measures with questions that aim to focus on the impairment caused by test anxiety. The AAT has a debilitating and a facilitating test anxiety scale, both of which were completed by participants in this study. Only the data from the debilitating test anxiety scale will be included within the thesis.

Procedure

After students signed-up for the study, the author met individually with each participant via Microsoft Teams. Informed consent to participate in the study, record the meeting, and include anonymous quotes in theses, publications, and presentations was obtained. The interviewer then proceeded with the study. First, the AAT as well as a demographic questionnaire that included questions about age, gender, current university term, and major/minor were administered to participants via Qualtrics. Participants completed the questionnaires with their video and/or audio turned off, without assistance from the interviewer aside from basic inquiries or technical difficulties. The interviewer was blind to AAT responses and scores. Second, the author administered the Test Anxiety Interview. The entirety of the study took 75 minutes on average. Given that the interview was semi-structured, it sometimes took

slightly more or less time to complete. All participants were renumerated with at least 1.25 SONA credits that could be used for bonus marks in applicable psychology courses. For those whose interviews took slightly longer, an extra 0.25 credits were rewarded for each additional 15 minutes.

Following completion of the study, the author assigned each participant a test anxiety severity score in order to separate participants into low, moderate, and high test anxiety groups. A rating system was adapted from the Anxiety and Related Disorders Interview Schedule for DSM-5 (ADIS-5) in order to remain as objective as possible (Brown & Barlow, 2014). The Diagnostic Summary Sheet on the ADIS-5 is typically used by clinicians to assign a clinical severity rating (CSR) to a diagnosis on the basis of the distress and interference that it appears to cause an individual. The severity scale ranges from 0 to 8. A rating of 0 means the disorder is not disturbing or disabling, whereas a rating of 8 means the disorder is very disturbing or disabling. Researchers have used this rating system to distinguish those with subclinical and clinical levels of various disorders by dividing the severity scale, whereby ratings of 0-3 denote subclinical diagnoses and ratings of 4-8 denote clinical diagnoses (Moscovitch et al., 2015). The author and her advisor chose to adopt a similar method in order to distinguish the three test anxiety groups in this study. It is important to note that the 9 point scale was adapted into an 11 point scale. This decision was made for a few reasons. First, we did not anticipate that ratings on the low and high ends of the scale would be assigned in this context. Ratings of 0 and 1 would apply to individuals whose test anxiety does not exist and/or does not cause any distress or impairment. In contrast, ratings of 7 or 8 would apply to individuals whose test anxiety has become so interfering that they would likely not be enrolled in university and thus would not be completing the study. Extending the scale thereby allowed us to capture nuance in the middle, where we anticipated the

majority of students in our sample would fall, while maintaining the integrity of the scale. Second, the rating system would be used to distinguish three groups (low, moderate, and high) in this study instead of two groups (clinical and subclinical) like in other studies (Moscovitch et al., 2015). Given that test anxiety is not an established disorder with criteria to guide decision-making, it seemed inappropriate to make a subclinical vs. clinical distinction. We also could not anticipate the range in test anxiety severity with which students who signed-up for the study would present. Separating participants into only two groups would have resulted in very unequal group sizes, which likely would have been misleading. Adapting the scale therefore made the distinction of three groups more feasible. For this study, ratings of 0-3 indicate low test anxiety, ratings of 4-5 indicate moderate test anxiety, and ratings of 6-10 indicate high test anxiety.

The author used her clinical judgement to formally assign these ratings based the distress and impairment test anxiety was reported to cause in the relevant sections of the interview once data collection had been completed. This was done to allow for a better understanding of test anxiety and to facilitate comparison, both of which would result in more accurate ratings. The author reviewed the interview notes and listened to the relevant portions of the recordings at this time. Given that the ratings are based on overall distress and interference caused by test anxiety, information from the fifth and sixth sections of the interview were primarily used to make these decisions. Information collected in other sections of the interview was not explicitly considered. Recall that the author was also blind to AAT responses and scores, so as not to influence decision-making. The author reviewed a subset of the ratings with her advisor to establish benchmarks before continuing with rating alone, given that it is a time-intensive task.

The author used a qualitative approach similar to that described by Gale et al. (2013). MS Teams automatically generated transcripts of the interviews. These transcripts were reviewed and

updated as needed by the author and undergraduate research assistants. The sections of the transcript that would be qualitatively analyzed were reviewed a second time by the author to ensure their accuracy. After becoming familiar with the data, the author completed content analyses of participants' responses to inquiries about their thoughts, emotions, and bodily reactions that were obtained in the second section of the interview. Nine content analyses were completed in total because each of these areas were broken down across time in relation to the test: thoughts (before, during, after the test), emotions (before, during, after the test), and bodily reactions (before, during, after the test). For these analyses, the author identified the important aspects of each participant's responses (e.g., only specific words that describe emotions for the content analyses pertaining to emotions). The author then grouped similar responses (e.g., anxious and nervous) together across participants. The content analyses resulted in a number of themes in each area, with which the author's advisor agreed. Labels were assigned to each theme and a coding manual that included examples was created. The relevant sections of the transcripts were then coded by the author and two other research assistants according to the manual. One participant was coded together in full while following the manual and discussing uncertainties. All of the other interviews were coded independently. After coding was complete, the author met with the research assistants to discuss the themes. At this time, some themes were removed due to low endorsement and/or ambiguity, while other themes were collapsed due to similarity. Reliability was calculated once the themes were finalized.

Reliability

Kappa values were used to estimate inter-rater reliability between the author and each of the research assistants. Inter-rater reliability was not calculated between the research assistants because it was decided beforehand that the author's codes would ultimately be those reported.

Thus, two kappa values were calculated for each theme within the nine content areas outlined previously. Given that several dozen themes exist overall, average kappa values were calculated. Two average kappa values will be reported for each of the nine areas, with the first value representing the agreement between the author and one research assistant and the second value representing the agreement between the author and the other research assistant. As we were striving for a strong level of agreement, average kappa values of .80 or above were considered acceptable.

Reliability for thoughts before the test was strong between the author and each coder, with average kappa values of 0.87 and 0.81. A higher level of disagreement was found with thoughts during and after the test, however. For thoughts during the test, the kappa value between the author and the first assistant did not fall within the acceptable range (0.77), whereas the kappa between the author and the second assistant did (0.80). The opposite was found for thoughts after the test: the kappa value between the author and the first research assistant fell within the acceptable range (0.89), whereas the kappa between the author and the second assistant did not (0.77). The three coders discussed disagreements within both of these areas, and the author's codes were updated with the final decisions.

In terms of emotions before the test, the level of agreement between the author and the first assistant was strong, with an average kappa value of 0.95. However, the average kappa value for the author and the second assistant fell outside of the acceptable range, with an average kappa of 0.78. The disagreements were discussed between coders, and the author's codes were updated with the final decisions. Reliability was strong between the author and both coders for emotions during and after the test. For emotions during the test, the average kappa values were 0.96 and 0.85. For emotions after the test, the kappa values were 0.95 and 0.84.

All of the average kappa values for bodily reactions fell within the acceptable range, indicating a high level of agreement between the author and both assistants. The average kappa values were 0.94 and 0.87 for bodily reactions before the test, 0.82 and 0.80 for bodily reactions during the test, and 0.90 and 0.90 for bodily reactions after the test. Given that reliability was strong, disagreements in this area were not reviewed.

Results

Sample Characteristics

Of the 36 undergraduate students who composed the final sample, 13 were assigned to the low group, 18 were assigned to the moderate group, and five were assigned to the high group. The low group was composed of 10 women and 3 men, and their mean age was 20 (SD = 1.22). Those in the low group were assigned a CSR of 2.62 on average (SD = .65, Range = 1 – 3). The moderate group was comprised of 14 women, 3 men, and 1 individual who identified as genderfluid, with an average age of 19.33 (SD = 1.28). The moderate group's mean CSR was 4.33 (SD = .49, Range = 4 – 5). Finally, the high group was made up of 3 women and 2 men, and the mean age was 20 (SD = 1.41). Participants in the high group had a CSR of 7.2 on average (SD = .84, Range = 6 – 8).

Test Anxiety Interview

The primary objective of this study was to identify the ways in which those with low, moderate, and high levels of test anxiety differ in terms of their thoughts, emotions, and bodily reactions within the context of testing situations. Participants were asked to recall all of the thoughts, emotions, and bodily reactions they were experiencing before, during, and after a recent test. The themes that emerged in each of these areas were compared among groups in order to determine whether differences exist based on the author's qualitative impressions. It is important to note that the differences were not tested.

Thoughts Before the Test

The author's content analysis for thoughts before the test resulted in 10 themes:

1. the grade that would be achieved on the test;
2. the importance of the test (e.g., for GPA, grad school, career);

3. how individuals believed they would perform during the test;
4. what individuals imagined the test would be like (e.g., difficult);
5. thoughts intended to calm or reassure oneself;
6. doubt in oneself (e.g., studying method, abilities);
7. memories of previous tests;
8. having too much to do before the test;
9. regret about studying; and
10. the test being over.

The two most popular themes across test anxiety groups were thoughts about the anticipated grade one would achieve and the importance of the test. Other themes that were commonly endorsed include thoughts about how one might perform on the test or what the test might be like, reflections of previous tests, self-doubt, and reassurance. Thinking about being done with the test was endorsed by a few participants in the low and moderate groups, but it was not endorsed by anyone in the high group. Thoughts about having too much to do before the test were reported by over one-third of participants in the low group but by only one participant in the moderate group and no participants in the high group. One student in the low group reported thinking, “I'm not going to be able to read all of this in time or go over all of this in time.” Another stated, “I tend to think about the other things that are due as well, especially when I have something more important due.” Table 1 shows the results in full.

Table 1

Number and Percentage of Participants who Endorsed Themes Under Thoughts Before the Test

	Low Group	Moderate Group	High Group
Grade	8 (61.54%)	14 (77.78%)	4 (80.00%)

Importance	6 (46.15%)	12 (66.67%)	5 (100.00%)
Performance	4 (30.77%)	6 (33.33%)	2 (40.00%)
Test	4 (30.77%)	4 (22.22%)	1 (20.00%)
Calm/Reassure	4 (30.77%)	5 (27.78%)	1 (20.00%)
Doubt	2 (15.38%)	8 (44.44%)	1 (20.00%)
Past Tests	3 (23.08%)	2 (11.11%)	3 (60.00%)
Too Much	5 (38.46%)	1 (5.56%)	0 (0.00%)
Regret Studying	2 (15.38%)	1 (5.56%)	1 (20.00%)
Being Done	3 (23.08%)	3 (16.67%)	0 (0.00%)

Thoughts During the Test

Seven themes emerged from the content analysis for thoughts during the test. Some of these themes were the same as those that emerged for thoughts before the test but others differed. The seven themes included:

1. the grade that would be achieved on the test;
2. the importance of the test (e.g., for GPA, grad school, career);
3. the time limit;
4. thoughts intended to calm or reassure oneself;
5. doubt in oneself (e.g., studying method, abilities);
6. questions that one does not know how to answer; and
7. the test being over.

Participants in each of the test anxiety groups commonly endorsed thinking about the time limit and the grade that would be achieved on the test. Self-doubt and a fixation on questions to which one did not know the answer were also frequently reported across groups. Although nearly half of the students in the low group described thinking about the importance of the test before writing it, this was not the case during its administration. None of the participants in the low group endorsed this theme, unlike the moderate and high groups. For example, one of the students in the high group remembered thinking, "how I am [doing] right now could dictate how I do in this course, what my grades look like, how it impacts my psych major". In addition to the anxious thoughts outlined, several participants in the low and moderate groups described having thoughts that were intended to calm and/or reassure themselves. A participant in the moderate group recalled having an internal, "pep talk to calm down." Another participant in the low group reported affirmations like, "it's OK" and, "you got this." In contrast, none of the students in the high group reported having these encouraging thoughts that seemingly countered the anxious thoughts. Furthermore, one of the participants in the high group could not recall having thoughts during the test at all, stating, "I don't know if there were any coherent thoughts going through my mind at the time. I think it just felt like a jumble, really, of just flashing red panic." See Table 2 for the number and percentage of participants in each group who endorsed the themes.

Table 2

Number and Percentage of Participants who Endorsed Themes Under Thoughts During the Test

	Low Group	Moderate Group	High Group
Grade	4 (30.77%)	6 (33.33%)	2 (40.00%)
Importance	0 (0.00%)	4 (22.22%)	2 (40.00%)

Time	4 (30.77%)	6 (33.33%)	2 (40.00%)
Calm/Reassure	3 (23.08%)	7 (38.89%)	0 (0.00%)
Doubt	2 (15.38%)	5 (27.78%)	3 (60.00%)
Do Not Know	2 (15.38%)	5 (27.78%)	3 (60.00%)
Being Done	2 (15.38%)	1 (5.56%)	1 (20.00%)

Thoughts After the Test

The content analysis for thoughts after the test revealed 12 themes. Although there is some overlap between the themes at this point in time and the themes before and during the test, mostly new themes emerged. The 12 themes are as follows:

1. the grade that would be achieved on the test;
2. the importance of the test (e.g., for GPA, grad school, career);
3. which questions one answered correctly or incorrectly;
4. actions one can take now to mitigate test performance;
5. thoughts intended to calm or reassure oneself;
6. future tests;
7. comparing oneself to others;
8. thoughts aimed at accepting one's performance on the test;
9. self-deprecation;
10. regret about studying and/or test performance;
11. conviction about being able to have done better on the test; and
12. the test being over.

Similar to before the test, the anticipated grade one would achieve was the most endorsed theme across test anxiety groups after the test. Thoughts about the test's importance, wonders about the questions on the test one might have answered correctly or incorrectly, thoughts to calm and/or reassure oneself, and being done with the test were also commonly reported by students in each group. The remaining seven themes were endorsed at a lower frequency overall, indicating diversity in participants' reactions. For example, some students reported thoughts of acceptance with regard to their test performance (e.g., "OK, I gotta get over it - I can't do anything now") whereas others endorsed self-deprecating thoughts (e.g., "I'm not going to get anywhere in life"). The largest discrepancy between groups in terms of these remaining themes is with respect to thoughts about future tests. Three out of five students in the high group endorsed this theme compared to one participant in each of the low and moderate groups. For example, one participant said, "It's going to repeat until the end of the term since [the quizzes] are weekly. So, I knew it was probably going to come back, all those feelings." See Table 3 for the frequency with which each theme was reported across groups.

Table 3

Number and Percentage of Participants who Endorsed Themes Under Thoughts After the Test

	Low Group	Moderate Group	High Group
Grade	8 (61.54%)	14 (77.78%)	3 (60.00%)
Importance	6 (46.15%)	5 (27.78%)	2 (40.00%)
Right/Wrong	3 (23.08%)	6 (33.33%)	1 (20.00%)
What Now	1 (7.69%)	3 (16.67%)	1 (20.00%)
Calm/Reassure	2 (15.38%)	5 (27.78%)	2 (40.00%)

Future Tests	1 (7.69%)	1 (5.56%)	3 (60.00%)
Comparison	2 (15.38%)	1 (5.56%)	2 (40.00%)
Acceptance	2 (15.38%)	2 (11.11%)	1 (20.00%)
Self-Deprecation	1 (7.69%)	1 (5.56%)	2 (40.00%)
Regret Studying	1 (7.69%)	4 (22.22%)	1 (20.00%)
Done Better	1 (7.69%)	3 (16.67%)	0 (0.00%)
Being Done	2 (15.38%)	6 (33.33%)	1 (20.00%)

Emotions Before the Test

Eight themes resulted from the content analysis for thoughts during the test. These include: guilt, shame, fear, worry, panic, anger, frustration, and sadness. Worry (e.g., about the test itself, grades) was the most common emotion among all test anxiety groups. Other emotions related to anxiety (i.e., fear and panic) were also frequently reported, along with feelings of frustration (e.g., with the material, at oneself) and guilt. Anger and shame, which may respectively be viewed as similar but more intense emotions to guilt and frustration, were less common but reveal differences between groups. First, none of the participants in the low group reported experiencing anger, whereas a percentage of those in the moderate and high groups did. Second, all but one of the participants in the high group communicated feeling shame, compared to very few in the low and moderate groups. One participant in the high group described their shame by expressing, “I felt I was wasting everybody's time and money by going into school if I was just going to fail.” Another recalled, “I did feel the shame of being a French immersion student who can barely study for a French test.” Finally, sadness was endorsed at the same rate as anger and guilt, though it was more evenly reported across groups. The results can be found in Table 4.

Table 4

Number and Percentage of Participants who Endorsed Themes Under Emotions Before the Test

	Low Group	Moderate Group	High Group
Guilt	6 (46.15%)	10 (55.56%)	5 (100.00%)
Shame	1 (7.69%)	2 (11.11%)	4 (80.00%)
Fear	8 (61.54%)	14 (77.78%)	2 (40.00%)
Worry	12 (92.31%)	18 (100.00%)	5 (100.00%)
Panic	4 (30.77%)	9 (50.00%)	3 (60.00%)
Anger	0 (0.00%)	5 (27.78%)	2 (40.00%)
Frustration	7 (53.85%)	9 (50.00%)	1 (20.00%)
Sadness	3 (23.08%)	3 (16.67%)	1 (20.00%)

Emotions During the Test

The majority of the same emotions that emerged before the test also emerged during. The first seven themes outlined before the test remained. The only thematic difference is that the eighth and final theme of sadness was replaced with relief during the test. The most frequently recalled emotions across the test anxiety groups were worry, fear, and panic. Like before the test, frustration (e.g., with the test, oneself) and guilt were also commonly reported themes overall, as were their similar but more intense counterparts (i.e., anger and shame). Fewer participants in the high group reported experiencing guilt and shame compared to nearly all participants in this group before the test. None of the participants in the high group described feeling angry during test administration, and only one described feeling frustrated. Relief was endorsed by at least one

student in each group. One participant described feeling relieved upon initially reviewing the test, recalling, “I don't think [the test] was as bad as the anxiety would have made it out to be.” Another explained that relief came up intermittently, “whenever I’d get something that I knew for sure.” This was the only positively-valanced emotional theme that emerged during or before the test. Table 5 shows these data.

Table 5

Number and Percentage of Participants who Endorsed Themes Under Emotions During the Test

	Low Group	Moderate Group	High Group
Guilt	4 (30.77%)	4 (22.22%)	2 (40.00%)
Shame	1 (7.69%)	2 (11.11%)	2 (40.00%)
Fear	4 (30.77%)	11 (61.11%)	3 (60.00%)
Worry	7 (53.85%)	11 (61.11%)	4 (80.00%)
Panic	2 (15.38%)	14 (77.78%)	4 (80.00%)
Anger	2 (15.38%)	5 (27.78%)	0 (0.00%)
Frustration	7 (53.85%)	9 (50.00%)	1 (20.00%)
Relief	1 (7.69%)	2 (11.11%)	1 (20.00%)

Emotions After the Test

In terms of emotions after the test, 10 themes were discovered. All of the emotional themes revealed before and during the test, including both sadness and relief, remained. Disappointment also emerged as a new theme. Worry (e.g., about grades, the future) persisted as the most prevalent theme across the test anxiety groups. The associated emotions of fear and

panic were less common compared to before and during the test. Guilt emerged as the second most popular theme among all groups (e.g., “I could have studied more”). Frustration and relief were also frequently endorsed. The percentage of those in the low group who endorsed feeling shame increased considerably after the test, whereas it remained similar for the moderate and high groups. Only one participant in each group reported feeling angry at this point in time. Sadness was endorsed by some participants in both the low and moderate groups but by none of the participants in the high group. The related yet more intense emotion of disappointment was favoured by the high group instead. The opposite was true for the low group, as none of these participants reported feeling disappointed afterwards. See Table 6 for the results in full.

Table 6

Number and Percentage of Participants who Endorsed Themes Under Emotions After the Test

	Low Group	Moderate Group	High Group
Guilt	5 (38.46%)	9 (50.00%)	3 (60.00%)
Shame	4 (30.77%)	1 (5.56%)	2 (40.00%)
Fear	4 (30.77%)	8 (44.44%)	2 (40.00%)
Worry	10 (76.92%)	12 (66.67%)	4 (80.00%)
Panic	1 (7.69%)	2 (11.11%)	1 (20.00%)
Anger	1 (7.69%)	1 (5.56%)	1 (20.00%)
Frustration	4 (30.77%)	6 (33.33%)	1 (20.00%)
Sadness	2 (15.38%)	4 (22.22%)	0 (0.00%)
Relief	3 (23.08%)	6 (33.33%)	2 (40.00%)

Disappointment	0 (0.00%)	2 (11.11%)	2 (40.00%)

Bodily Reactions Before the Test

The following 11 bodily reactions were identified as themes before the test: muscle tension, change in heart rate (e.g., increased, irregular), change in temperature (e.g., too hot and/or cold, deregulated), change in breathing (e.g., increased, difficulty), sweating, shaking, headache, stomachache, nausea, dizziness, tight chest. Experiencing a change in heart rate was the most frequently reported bodily reaction across test anxiety groups before the test. This was followed closely by muscle tension. A change in breathing was common in the moderate and high groups but not as prevalent in the low group, whereas a change in temperature was common in the low and moderate groups but not as widespread in the high group. Stomachaches, headaches, and sweating were also moderately endorsed. Experiencing a tight chest and dizziness were less frequently reported in each group but are still notable. The final bodily reaction of nausea was described by three out of five students in the high group, yet only one student in moderate group and no participants in the low group. Therefore, quite a bit of diversity in bodily reactions was demonstrated by participants. Table 7 shows the number and percentage of participants who endorsed each theme across groups.

Table 7

Number and Percentage of Participants who Endorsed Themes Under Bodily Reactions Before the Test

	Low Group	Moderate Group	High Group
Muscle Tension	6 (46.15%)	10 (55.56%)	3 (60.00%)

Change in Heart Rate	7 (53.85%)	14 (77.78%)	3 (60.00%)
Change in Breathing	2 (15.38%)	9 (50.00%)	4 (80.00%)
Change in Temperature	4 (30.77%)	9 (50.00%)	1 (20.00%)
Sweating	1 (7.69%)	10 (55.56%)	1 (20.00%)
Shaking	2 (15.38%)	10 (55.56%)	2 (40.00%)
Headache	2 (15.38%)	6 (33.33%)	3 (60.00%)
Stomachache	4 (30.77%)	6 (33.33%)	3 (60.00%)
Nausea	0 (0.00%)	1 (5.56%)	3 (60.00%)
Dizziness	1 (7.69%)	3 (16.67%)	1 (20.00%)
Tight Chest	1 (7.69%)	3 (16.67%)	2 (40.00%)

Bodily Reactions During the Test

Most of the same bodily reactions that emerged as themes before the test also emerged during. The first 10 themes outlined before the test remained unchanged. In fact, the only thematic difference is that the eleventh and final theme of a tight chest was replaced with restlessness. A change in heart rate was again the most frequently reported theme across groups. Muscle tension, sweating, and a change in temperature were also commonly shared. Shaking and a change in breathing were typical of both the moderate and high groups, but not necessarily of the low group. In addition, headaches were reported by three participants in both the moderate and high groups but by no one in the low group. Stomachaches, dizziness, and nausea were endorsed by five or fewer students overall, indicating continued diversity in participants' physiological reactions. Three out of five students in the high group reported feeling nauseous

during test administration, but this theme did not emerge in the low or moderate groups. See Table 8 for the results in full.

Table 8

Number and Percentage of Participants who Endorsed Themes Under Bodily Reactions During the Test

	Low Group	Moderate Group	High Group
Muscle Tension	4 (30.77%)	6 (33.33%)	3 (60.00%)
Change in Heart Rate	4 (30.77%)	15 (83.33%)	3 (60.00%)
Change in Breathing	2 (15.38%)	4 (22.22%)	4 (80.00%)
Change in Temperature	3 (23.08%)	5 (27.78%)	4 (80.00%)
Sweating	3 (23.08%)	8 (44.44%)	3 (60.00%)
Shaking	1 (7.69%)	11 (61.11%)	4 (80.00%)
Headache	0 (0.00%)	3 (16.67%)	3 (60.00%)
Stomachache	0 (0.00%)	4 (22.22%)	1 (20.00%)
Nausea	0 (0.00%)	0 (0.00%)	3 (60.00%)
Dizziness	1 (7.69%)	2 (11.11%)	1 (20.00%)
Restlessness	3 (23.08%)	3 (16.67%)	1 (20.00%)

Bodily Reactions After the Test

Fewer bodily reactions were discovered as themes after the test. Each of the nine themes were duplicates of those that emerged before and/or during. Restlessness, dizziness, and nausea

were the only themes from the other points in time that were not apparent after the test. Overall, bodily symptoms were reported at a much lower rate by those in low group, with two or fewer participants out of over a dozen reporting each theme. They persisted more commonly, but still at a lower rate compared to before and during the test, for those in the moderate and high groups. Muscle tension, a change in breathing, and headaches were most frequently endorsed. A change in heart rate remained common in the moderate group, but none of the participants in the high group reported this. The remaining themes (i.e., stomachache, shaking, a change in temperature, sweating, and tight chest) were endorsed less frequently overall, once again demonstrating diversity in terms of the body's anxiety response. These results can be found in Table 9.

Table 9

Number and Percentage of Participants who Endorsed Themes Under Bodily Reactions After the Test

	Low Group	Moderate Group	High Group
Muscle Tension	2 (15.38%)	4 (22.22%)	3 (60.00%)
Change in Heart Rate	1 (7.69%)	5 (27.78%)	0 (0.00%)
Change in Breathing	0 (0.00%)	5 (27.78%)	3 (60.00%)
Change in Temperature	0 (0.00%)	2 (11.11%)	2 (40.00%)
Sweating	1 (7.69%)	2 (11.11%)	1 (20.00%)
Shaking	1 (7.69%)	2 (11.11%)	1 (20.00%)
Headache	1 (7.69%)	5 (27.78%)	2 (40.00%)
Stomachache	1 (7.69%)	2 (11.11%)	2 (40.00%)

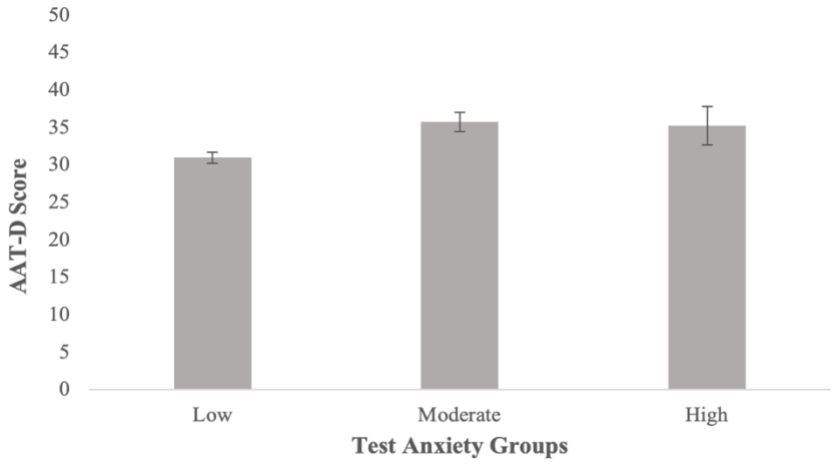
Tight Chest	0 (0.00%)	3 (16.67%)	1 (20.00%)

Test Anxiety Questionnaire

The secondary objective for this study was to evaluate one of the existing test anxiety questionnaires (i.e., AAT). A one-way ANOVA (two-tailed, $\alpha = .05$) was conducted to determine whether the test anxiety groups differed from one another in their AAT-D scores. The one-way ANOVA showed that there was a statistically significant difference between groups ($F(2,33) = 4.16$, $p = .024$). A Tukey post-hoc test found that those in the low group ($M = 30.92$, $SD = 2.72$) scored significantly lower than the moderate group on the questionnaire ($M = 35.72$, $SD = 5.44$, $p = .022$). A statistically significant difference was not found between the low and high groups ($M = 35.20$, $SD = 5.81$, $p = .209$), nor the moderate and high groups ($p = .974$). It is important to note that the effect size of the low and moderate group comparison is similar to the low and high group comparison; however, the latter was not significant due to the smaller sample size. See Figure 1 for a bar graph depicting the average AAT-D scores among groups.

Figure 1

Means and Standard Errors of AAT-D Questionnaire Scores for the Low, Moderate, and High Test Anxiety Groups. Error Bars Represent the Standard Error for Each Group.



Discussion

The main objective of this study was to identify the ways in which the experience of test anxiety qualitatively differs in those with objectively low, moderate, and high levels of test anxiety. Specifically, the thoughts, emotions, and bodily reactions that characterize this experience were investigated through interviews. A secondary aim was to determine whether the test anxiety groups differ in terms of their scores on a test anxiety questionnaire. Analyses of students' responses to relevant inquiries on the interview and their questionnaire scores help to answer the identified research questions, contribute to the growing body of literature on test anxiety, and advance our understanding of how best to conceptualize, measure, and treat this phenomenon.

Before analyzing the differences that emerged through the interview, it is important to highlight that many similarities were found in the thoughts, emotions, and bodily reactions individuals with different levels of test anxiety experience. For example, a large number of students across groups reported having negative thoughts (e.g., fixating on their grade), feeling worried, fearful, and guilty, and experiencing a variety of bodily reactions, such as a change in heart rate and muscle tension. These findings are in line with our understanding of some established anxiety and anxiety-related disorders. What distinguishes those with subclinical and clinical anxiety is not necessarily the symptoms that individuals experience, but rather the excessiveness of those symptoms to the situation at hand as well as the distress and/or interference they cause (American Psychiatric Association, 2013). Naturally, however, there is diversity in individuals' cognitive, emotional, and bodily reactions to anxiety. Although many similarities were found, the experiences of the students in this study were not homogenous

within or between groups. Some of the main qualitative differences that were demonstrated among the test anxiety groups will now be discussed.

Thoughts

Before the test, over one-third of participants in the low test anxiety group endorsed thoughts about having too much to do in a short amount of time; however, this theme was reported by only one participant in the moderate group, and it was not reported by any participants in the high group. It may be reasonably assumed that a student would have thoughts like this most often in two situations. First, these thoughts may arise if a student has an upcoming test scheduled for a particularly busy point in the term (i.e., the midterm or exam period). Second, these thoughts may surface if a student began studying for an upcoming test at a later point than required to comprehensively review the material. In either of these situations, it would make sense for a student to feel anxious as a result of being overwhelmed or unprepared, respectively. Perhaps then at least some students in the low group experience test anxiety more sensibly in the context of an objective stressor, rather than in ordinary circumstances without an identifiable stressor like those in the moderate and high groups might. It is possible that for students in the moderate and/or high group, testing in and of itself is the stressor that brings on feelings of anxiety.

Two interesting differences in thought were revealed during the test. First, none of the participants in the low group endorsed worrying about the importance of the test (e.g., final grade, GPA, grad school) during its administration, whereas those in the moderate and high groups did. Interestingly, this difference was not found before the test, as thinking about the test's importance was common across groups. It seems as though all of the participants in the low group who were previously thinking about the stakes of the test were able to block out these

thoughts while writing it, likely in order to focus on the test's completion. Although some participants in the moderate and high groups seem to have also been able to accomplish this, as evidenced by less endorsement compared to before the test, these thoughts persisted for some students. The second difference that was found during the test may reveal an important distinction between groups in terms of coping ability. As previously mentioned, several anxious thoughts (e.g., self-doubt, fixation on the time limit and questions that one did not know how to answer) were reported across groups. Seemingly to de-escalate their anxiety, those in the low and moderate groups reported having thoughts to calm and/or reassure themselves. In contrast, none of the participants in the high group reported these counter thoughts that may help to regulate anxiety. For this group, anxiety during the test may go unchecked and continue to intensify, which might explain why it becomes interfering and causes distress.

Due to the diversity in cognitions after the test, no clear differences were identified among the groups. The largest discrepancy emerged regarding thoughts about future tests. Three of the five participants in the high group already started worrying and/or planning for another upcoming test. In other words, they moved onto the next time the feared situation would arise immediately. In contrast, only one participant in each of the low and moderate groups endorsed this theme. This may suggest that these groups are more readily able to move on from the most recent testing situation without fear of the next; however, this difference must be interpreted with caution.

Emotions

Before the test, four of the five participants in the high test anxiety group reported feeling shame, compared to only one participant in the low group and two in the moderate group. Shame in the high group was mostly attributed to the difficulty associated with testing and/or feeling

like an outsider compared to other students because of it. It is possible that shame was not as highly endorsed in the other groups because their struggles with testing are not as severe and thus feel more comparable to that of other students. This result is consistent with Pekrun et al.'s (2004) findings that anxiety and shame are positively correlated with one another in the context of test-taking. Many participants in the low and moderate groups endorsed feeling frustrated and/or angry during the test. In contrast, only one participant in the high group reported frustration and none of the participants reported anger. In the low and moderate groups, these feelings were directed at the test and/or themselves. Perhaps these emotions were not endorsed as often in the high group because they were outweighed by feelings associated with anxiety (i.e., worry, fear, panic) which were more intense; however, this is mostly speculation. No obvious differences emerged among the test anxiety groups in terms of emotions after the test.

Bodily Reactions

The diversity in the combination of bodily symptoms experienced by participants makes it difficult to meaningfully interpret the differences that were demonstrated among the groups. The one difference that stands out before and during the test is that nausea was experienced almost exclusively by those in the high group. Three participants in the high group reported feeling nauseous before and during test administration, compared to only one participant in the moderate group before and none of the participants in either group during. Nausea is one of, if not the most, severe bodily symptom that was reported at any point. This reaction to anxiety would undoubtedly make it very difficult to concentrate on studying and/or writing an exam, which may contribute to why test anxiety for those in the high group is so interfering. After the test, few bodily reactions were reported by participants in the low group, suggesting that anxiety decreased quickly. The reactions persisted at a higher rate for participants in the moderate and

high groups, indicating a more prolonged anxiety response and/or a slower recovery. It is important to note that some participants' bodily reactions lasted for an extended period of time following the test. For example, one participant in the high group reported, "I still had that migraine. I actually had to go to the hospital for that migraine because it would not go away." This highlights that test anxiety is not only a problem before and during tests but that its effects may continue to interfere with academic and non-academic tasks well beyond the completion of a test, further adding to the interference and distress it can cause for some individuals.

Questionnaire

The questionnaire results indicate that the AAT-D is not necessarily able to distinguish test anxiety severity. A significant difference was detected between the low and moderate groups; however, neither of the other comparisons were not found to be significant. Although a significant difference may have been uncovered between the low and high groups with a larger sample size, the moderate and high groups' mean scores were very similar to one another. In fact, some of the participants in the moderate group received a greater score than those in the high group, which is certainly a problem. Given that it is likely the distinction between these two groups that causes the most confusion, it is imperative that a measure is able to make this distinction. However, the results suggest that the AAT-D does not differentiate between those whose test anxiety is a problem (i.e., high group) and those whose test anxiety is not (i.e., low and moderate groups). Although this is the most straightforward explanation, these results may also be due to an issue with the groupings. As such, more research in this area is required.

Implications

The number of similarities that exist between the test anxiety groups in this study substantiate the notion that current measures of test anxiety include questions that are likely

applicable to a wide range of students. For example, the majority of the existing questionnaires (e.g., AAT, CTAS, RTA, RTT, TEQ, WTAS) enquire about whether respondents worry about grades. Recall that nearly all of the participants in this study reported thinking about their grades before, during, and after the test. Therefore, questions of this nature would likely be answered similarly regardless of test anxiety severity. In addition to this, current test anxiety measures also seem to include questions that may not be applicable to those with high levels of test anxiety. For example, most of the existing questionnaires ask about an array of physiological reactions to anxiety in the context of testing. The current study found that there is substantial diversity in the combination of bodily reactions that students may exhibit. As such, asking about individual bodily reactions (e.g., sweating, shaking, change in temperature) does not seem to be a feasible way of assessing test anxiety severity. It is possible that those who belong in a low or moderate group might endorse the bodily reactions that are asked about on any given measure, whereas those who belong in a high group may find those particular reactions to be inapplicable to their experience. This may result in individuals in the former groups obtaining higher scores than those in the latter group, which did occur in some cases in the current study. This is certainly problematic, given that researchers often impose cut-off scores on questionnaires to distinguish test anxiety severity groups.

The results of the current study therefore offer evidence for the measurement issues that have been identified within the field. The findings suggest that existing questionnaires include potentially unsuitable and/or ineffective questions that do not necessarily reflect differences among those with varying test anxiety severity levels. This offers support for the notion that it is inappropriate to utilize cut-offs on these measures to make a distinction regarding test anxiety severity. This was shown to be the case with the AAT-D, as a significant difference was found

only between the low and moderate groups, and it is likely also applicable to other measures given that many are similar in nature. As such, these results could be used as a starting point to develop a refined measure to assess test anxiety severity based on the differences found. This would facilitate future research on the conceptualization of test anxiety, as much is still unknown in terms of how it fits within the context of other types of anxiety. It would also make identifying students with high levels of test anxiety feasible for accessibility offices and counsellors should universities be willing to offer academic accommodations in the future. A final implication of this research is that these findings may highlight the most beneficial intervention targets to explore. For example, it seems that self-regulation during the test is important to address, given that those in the high test anxiety group did not recall having any calming or reassuring thoughts to counteract their anxious or negative cognitions. In sum, these results have the potential to be influential both theoretically and clinically.

Limitations and Future Directions

The next step in this line of research would likely involve a follow-up study to confirm the differences that were found and to explore others. Ideally a follow-up study would address some of the current study's limitations. First and foremost are the unequal group sizes. Given that the author used a novel method to separate participants into groups, it was nearly impossible to plan for an even distribution. This is especially true given that the participants were only assigned to groups after the interviews were conducted. The inequality made it difficult to interpret the results and evaluate their validity in some circumstances. For a follow-up study, it would be helpful to briefly screen potential participants with a subset of questions prior to completion of the full interview in an effort to ensure more equal group sizes.

Although we have referred to the novel method that was used to distinguish groups as objective, it certainly is not free from bias. This method seems to be more objective than the traditional approaches to assessing test anxiety severity (i.e., self-report measures); however, it still ultimately relied on the clinical judgement of one individual. Ideally, multiple raters would come to a consensus on each participants' severity rating and group placement in a future study. Longer term, a psychometrically sound and well-validated interview tool to identify those whose test anxiety is a problem would be useful in both research and clinical settings. This is perhaps the closest we can come to a truly objective assessment of test anxiety for the time being.

Another limitation is that the participants in this study were enrolled in various university programs and were at different stages in their academic careers. Both of these variables have the potential to impact the results. Perhaps students in certain majors are more likely to experience test anxiety than others. Varying levels of experience in university may also impact comfortability with tests. Furthermore, testing demands likely differ as a result of these characteristics which would also reasonably impact anxiety levels. For a more equitable comparison moving forward, a more uniform group of participants would be useful to reduce demographic differences among groups.

A final limitation is that the length of the interview did not allow the author to collect information about participants' struggles with other types of anxiety. It would have been extremely valuable to have this information to compare among the test anxiety groups. Inquires of this nature should be incorporated into future studies. Not only would this advance our understanding of the conceptualization of test anxiety, but it would thereby contribute to improving its measurement and optimizing its treatment.

References

Alpert, R., & Haber, R. N. (1960). Anxiety in academic achievement situations. *Journal of Abnormal and Social Psychology, 61*(2), 207–215. 10.1037/h0045464.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). American Psychiatric Association.

Beidel, D. C., & Turner, S. M. (1988). Comorbidity of test anxiety and other anxiety disorders in children. *Journal of Abnormal Child Psychology, 16*(3), 275–287. 10.1007/bf00913800

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*(2), 169–179. 10.1016/0887-6185(94)90014-0

Benson, J., & El-Zahhar, N. (1994). Further refinement and validation of the revised test anxiety scale. *Structural Equation Modeling, 1*(3), 203–221. 10.1080/10705519409539975

Bögels, S. M., Alden, L., Beidel, D. C., Clark, L. A., Pine, D. S., Stein, M. B., & Voncken, M. (2010). Social anxiety disorder: Questions and answers for the DSM-V. *Depression and Anxiety, 27*(2), 168–189. 10.1002/da.20670

Brown, C. H. (1938). Emotional reactions before examinations: II. Results of a questionnaire. *The Journal of Psychology, 5*(1), 11–26. 10.1080/00223980.1938.9917549

Brown, T. A., & Barlow, D. H. (2014). *Anxiety and Related Disorders Interview Schedule for DSM-5: Client interview schedule.* Oxford University Press.

Brown, T. A., Chorpita, B. F., Korotitsch, W., & Barlow, D. H. (1997). Psychometric properties of the Depression Anxiety Stress Scales (DASS) in clinical samples. *Behaviour Research and Therapy, 35*(1), 79–89. 10.1016/S0005-7967(96)00068-X

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

Chiang, B., & Purdon, C. (2019). *The voice of doubt in obsessive-compulsive disorder*. Manuscript in preparation.

Driscoll R. (2007). *Westside Test Anxiety Scale validation.* Retrieved from https://files.eric.ed.gov/fulltext/ED495968.pdf

Gale, N. K., Heath, G., Cameron, E., Rashid, S., & Redwood, S. (2013). Using the framework method for the analysis of qualitative data in multi-disciplinary health research. *BMC Medical Research Methodology, 13*(1), 117–124. 10.1186/1471-2288-13-117

Hembree, R. (1988). Correlates, causes, effects, and treatment of test anxiety. *Review of Educational Research, 58*(1), 47–77. 10.2307/1170348

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

Johanson, G. A., & Brooks, G. P. (2010). Initial scale development: Sample size for pilot studies. *Educational and Psychological Measurement, 70*(3), 394–400. 10.1177/0013164409355692

Lazarus, R. S., & Folkman, S. (1984). *Stress, appraisal, and coping.* Springer.

LeBeau, R. T., Glenn, D., Liao, B., Wittchen, H. U., Beesdo-Baum, K., Ollendick, T., & Craske, M. G. (2010). Specific phobia: A review of DSM-IV specific phobia and preliminary recommendations for DSM-V. *Depression and Anxiety, 4*(2), 148–167. 10.1002/da.20655

Liebert, R. M., & Morris, L. W. (1967). Cognitive and emotional components of test anxiety: A distinction and some initial data. *Psychological Reports, 20*(3), 975–978. 10.2466/pr0.1967.20.3.975

Lovibond, P. F., & Lovibond, S. H. (1995). The structure of negative emotional states: Comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. *Behaviour Research and Therapy, 33*(3), 335–343. 10.1016/0005-7967(94)00075-u

Lowe, P. A., Lee, S. W., Witteborg, K. M., Prichard, K. W., Luhr, M. E., Cullinan, C. M., Mildren, B. A., Raad, J. M., Cornelius, R. A., & Janik, M. (2008). The Test Anxiety Inventory for Children and Adolescents (TAICA): Examination of the psychometric properties of a new multidimensional measure of test anxiety among elementary and secondary school students. *Journal of Psychoeducational Assessment, 26(*3), 215–230. 10.1177/0734282907303760

Mandler, G., & Sarason, S. B. (1952). A study of anxiety and learning. *Journal of Abnormal Psychology, 47*(2), 166–173. 10.1037/h0062855

McDonald, A. S. (2001). The prevalence and effects of test anxiety in school children. *Educational Psychology*, *21*(1), 89–101. 10.1080/01443410020019867

Moscovitch, D. A., Shaughnessy, K., Waechter, S., Xu, M., Collaton, J., Nelson, A. L., Barber, K. C., Taylor, J., Chiang, B., & Purdon, C. (2015). A model for recruiting clinical research participants with anxiety disorders in the absence of service provision: Visions, challenges, and norms within a Canadian context. *The Journal of Nervous and Mental Disease*, *203*(12), 943–957. 10.1097/NMD.0000000000000400

Osterhouse, R. A. (1972). Desensitization and study-skills training as treatment for two types of test-anxious students. *Journal of Counseling Psychology, 19*(4), 301–307. 10.1037/h0034177

Pekrun, R., Goetz, T., Perry, R. P., Kramer, K., Hochstadt, M., & Molfenter, S. (2004). Beyond test anxiety: Development and validation of the test emotions questionnaire (TEQ). *Anxiety, Stress, and Coping, 17*(3), 287–316. 10.1080/10615800412331303847

Reiss, N., Warnecke, I., Tolgou, T., Krampen, D., Luke-Krausgrill, U., & Rohrmann, S. (2017). Effects of cognitive behavioral therapy with relaxation vs. imagery rescripting on test anxiety: A randomized controlled trial. *Journal of Affective Disorders, 208*, 483–489. 10.1016/j.jad.2016.10.039

Sarason, I. G. (1958). Interrelationships among individual difference variables, behavior in psychotherapy, and verbal conditioning. *Journal of Abnormal and Social Psychology, 56*(3), 339–344. 10.1037/h0048353

Sarason, I. G. (1972). Experimental approaches to test anxiety: Attention and the uses of information. In C. D. Spielberger (Ed.), *Anxiety: Current trends in theory and research* (Vol. 2, pp. 381–403). Academic Press.

Sarason, I. G. (1984). Stress, anxiety, and cognitive interference: Reactions to tests. *Journal of Personality and Social Psychology, 46*(4), 929–938. 10.1037/0022-3514.46.4.929

Segool, N. K., von der Embse, N. P., Mata, A. D., & Gallant, J. (2014). Cognitive behavioral model of test anxiety in a high-stakes context: An exploratory study. *School Mental Health, 6*(1), 50–61. 10.1007/s12310-013-9111-7

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). Academic Press.

Spielberger, C. D., Anton, W. D., & Bedell, J. (1976). The nature and treatment of test anxiety. In M. Zuckerman & C. D. Spielberger (Eds.), *Emotions and anxiety: New concepts, methods, and applications* (pp. 317–345). Psychology Press. 10.4324/9781315744643

Spielberger, C. D., Gonzalez, H. P., Taylor, C. J., Algaze, B., & Anton, W. D. (1978). Examination stress and test anxiety. In C. D. Spielberger & I. G Sarason (Eds.), *Stress and anxiety* (Vol. 5, pp. 167–191). Hemisphere.

Suinn, R. M. (1969). The STABS, a measure of test anxiety for behavior therapy: Normative data. *Behaviour Research and Therapy, 7*(3), 335–339. 10.1016/0005-7967(69)90018-7

Thomas, C. L., Cassady, J. C., & Holmes Finch, W. (2018). Identifying severity standards on the Cognitive Test Anxiety Scale: Cut score determination using latent class and cluster analysis. *Journal of Psychoeducational Assessment, 36*(5), 492–508. 10.1177/0734282916686004

von der Embse, N., Jester, D., Roy, D., Post, J. (2018). Test anxiety effects, predictors, and correlates: A 30-year meta- analytic review. *Journal of Affective Disorders, 227*, 483–493. 10.1016/j.jad.2017.11.048.

Zeidner, M. (2014). Test anxiety. In P. Emmelkamp & T. Ehring (Eds.), *The Wiley handbook of anxiety disorders* (1st ed., pp. 581–595). John Wiley & Sons, Ltd.

Zeidner, M., & Matthews, G. (2003). Test Anxiety. In R. Fernández-Ballesteros (Ed.), *Encyclopedia of psychological assessment* (Vol. 1., pp. 965-969). SAGE Publications Ltd.

Zuriff, G. E. (1997). Accommodations for test anxiety under ADA? *Journal of the American Academy of Psychiatry and the Law*, *25*(2), 197–206.

Appendix

Test Anxiety Interview

STUDY ID#:		SONA ID#:		DATE:	
GENDER:		AGE:			

Many people feel anxious about tests. Anxiety about tests can occur before, during, and/or after a test. When we are anxious we experience a range of emotional and physical responses and thinking patterns. We are interested in gaining a better understanding of people's experience of test anxiety.

For the purpose of this interview, the term *test* will be used to refer to tests, midterms, and/or exams, whatever their format (essay, multiple choice, short answer, bell ringer, oral, mid-term, final, etc.). We will use the term *test anxiety* to mean anxiety people feel about tests *before* the test (that is, the time between when the anxiety about an upcoming test starts and when you enter the room in which the test will occur), *during* the test (the time between when you enter the room in which it will occur until you have handed the test in or otherwise completed it), and *after* the test (the time between when you completed the test and the anxiety about it fades, including the time just before and after receiving your test result).

[Note: The researcher will take notes regarding participants' responses to interview questions on this form].

BACKGROUND INFORMATION ABOUT RECENT TEST ANXIETY EXPERIENCE

I'd like you to identify a test you completed within the last year at university that you were feeling especially anxious about, and then I am going to ask you some questions about that test.

Can you think of a recent test that you were feeling especially anxious about that would not be overwhelming for you to discuss? Tell me about this test.

[If the participant has more than one experience in mind, ask the participant to focus on the most recent experience].

[As the participant answers this question, fill in the answers to the questions below. If the participant does not provide some of this information in the response, ask the remaining questions below to fill in the missing information].

a. How long ago was the test? What term?

b. How far into the term was the test scheduled (e.g., midterm, final exam)?

    i. Was this the first test you had written for the course?

ii. Was this test scheduled for a particularly busy point in the semester when you also had other major tests or/and assessments scheduled?
iii. Was this your final assessment for the course?

c. What course was the test for?

i. Was this course within your major/minor?

ii. Was this a mandatory or elective course?

iii. Had you previously taken other courses similar in content to this course? If so, what was your experience in these previous courses?

d. What was your approximate grade in the course prior to the test?

e. What percentage of your grade was the test worth?

f. What was the format of the test (e.g., multiple choice, short answer, essay)?

g. How much time did you have to complete the test?

h. Did you receive any academic accommodations for the test?

i. What approximate grade did you achieve on the test?

j. What other assessments were used in this course (e.g., other tests, essays, assignments)?

i. What approximate grades did you achieve on these other assessments?

k. What approximate final grade did you achieve in the course?

EXPERIENCE OF TEST ANXIETY

I want you to continue thinking about this particular test as you answer the following questions. I would like to get some more information about your experience of anxiety about this test.

I would like you to walk me through your experience of anxiety about this test. Think of your anxiety as occurring along a timeline. Start by explaining when your anxiety about the test first began and continue describing how your anxiety progressed until it went away. I would like you to go back in time as if you are currently experiencing the anxiety. You can close your eyes to visualize this more clearly if you feel comfortable with that. I will give you a minute or so to collect your thoughts before getting started.

[Wait one minute].

You can get started whenever you are ready.

[If the participant does not clearly identify when their anxiety started (and what triggered it), ended, and peaked, clarify this with them afterwards. Also clarify whether the participant's anxiety faded after the test and if so, whether it was reactivated before and/or after learning the grade achieved on the test].

Now I am going to ask you some more specific questions about your experience.

Before the Test [Ask the following questions if the participant reported experiencing anxiety at any point before the test. If not, ask whether the participant experienced any anxiety before the test. If the participant's response is no, move onto "During the Test"].

I would like you to think about the period of time before the test. That is the time from when the anxiety about this test started until you entered the room in which the test occurred. While you were experiencing anxiety before the test …

1. What thoughts were going through your mind?

    a. Many people experience an internal monologue, where they have a running speech in their head that represents their thoughts. It is not at all uncommon for us to experience our thoughts as being "spoken" to us, as talking to ourselves, or as someone talking to us. Did any of the thoughts that were going through your mind appear in this way? If so, tell me about that.

    b. Were any of these thoughts experienced as images, memories, flashforwards, or predictions? If so, tell me about that.

2. What were all the emotions you were feeling? [Prompt: What specific words would you use to describe your feelings (e.g., anger, frustration, guilt, shame, worry, panic, fear]?

    a. Which emotion was the strongest before the test? How would you rate the intensity of this emotion on a scale of 0-100?

3. What bodily symptoms did you notice (e.g., racing heart, stomach-ache, difficulty breathing, sweating, shaking, flushing, cold hands or feet, headache, clenched jaw or shoulders, increased breathing, tingling)?

4. Were anxious thoughts, emotions, or bodily symptoms most prominent before the test?

During the Test [Ask the following questions if the participant reported experiencing anxiety at any point during the test. If not, ask whether the participant experienced any anxiety during the test. If the participant's response is no, move onto "After the Test"].

Now I would like you to think about the period of time during the test. That is the time from when you entered the room in which the test occurred until you handed the test in or otherwise completed it. While you were experiencing anxiety during the test …

5. What thoughts were going through your mind?

    a. Did you experience any of these thoughts as being "spoken" to you, as talking to yourself, or as someone talking to you? If so, tell me about that.

    b. Were any of these thoughts experienced as images, memories, flashforwards, or predictions? If so, tell me about that.

6. What were all the emotions you were you feeling? [Prompt: What specific words would you use to describe your feelings (e.g., anger, frustration, guilt, shame, worry, panic, fear]?

    a. Which emotion was the strongest during the test? How would you rate the intensity of this emotion on a scale of 0-100?

7. What bodily symptoms did you notice (e.g., racing heart, stomach-ache, difficulty breathing, sweating, shaking, flushing, cold hands or feet, headache, clenched jaw or shoulders, increased breathing, tingling)?

8. Were anxious thoughts, emotions, or bodily symptoms most prominent during the test?

<u>After the Test [Ask the following questions if the participant reported experiencing anxiety at any point after the test. If not, ask whether the participant experienced any anxiety after the test. If the participant's response is no, move onto "Perceived Demands of the Test and Ability to Cope"].</u>

Now I would like you to think about the period of time after the test. That is the time from when you completed the test until the anxiety about it faded. While you were experiencing anxiety after the test …

9. What thoughts were going through your mind?

    a. Did you experience any of these thoughts as being "spoken" to you, as talking to yourself, or as someone talking to you? If so, tell me about that.

    b. Were any of these thoughts experienced as images, memories, flashforwards, or predictions? If so, tell me about that.

10. What were all the emotions you were you feeling? [Prompt: What specific words would you use to describe your feelings (e.g., anger, frustration, guilt, shame, worry, panic, fear]?

    a. Which emotion was the strongest after the test? How would you rate the intensity of this emotion on a scale of 0-100?

11. What bodily symptoms did you notice (e.g., racing heart, stomach-ache, difficulty breathing, sweating, shaking, flushing, cold hands or feet, headache, clenched jaw or shoulders, increased breathing, tingling)?

12. Were anxious thoughts, emotions, or bodily symptoms most prominent after the test?

PERCEIVED DEMANDS OF THE TEST AND ABILITY TO COPE

I want you to answer the next few questions as if you have gone back in time to the point when you were experiencing the highest level of anxiety before you wrote the test. To orient you to this period of time, I want you to describe to me what you are picturing. How long before the test is this? Where are you? What are you feeling at this moment? I want you to respond to the following questions according to your beliefs at the time you are picturing right now before you wrote the test, not according to how you feel or think in the present. Does that make sense?

[If the participant did not report experiencing anxiety before the test, ask them to focus on the point at which they were feeling the most anxiety during the test. If the participant did not report experiencing anxiety before the test or during the test, ask them to focus on the point at which they were feeling the most anxiety after the test].

<u>Perceived Demands</u>

1. What are all the ways the test is going to be challenging (e.g., time limit, format)?

    a. How challenging do you think the test going to be on a scale of 0 (not challenging at all) to 10 (extremely challenging)?

0 1 2 3 4 5 6 7 8 9 10

2. What are all the ways the test is important (e.g., for your final mark, standing in the program, future, disappointing self, disappointing family)?

    a. How important is the test to you **overall** on a scale of 0 (not important at all) to 10 (extremely important)?

0 1 2 3 4 5 6 7 8 9 10

3. Is there anything else that is making you feel especially anxious about this test?

<u>Perceived Ability to Cope</u>

1. How do you feel about your ability to effectively study for the test (e.g., focus, retain information)?

2. How do you feel about your ability to effectively write the test?

3. How do you think your family or other important people in your life (e.g., family, friends) feel about your ability to effectively study and write the test in order to do well on it?

4. How do you feel about your ability to effectively manage your anxiety?

5. How confident are you in your ability to effectively study and write the test **overall** on a scale of 0 (not confident at all) to 10 (extremely confident)?

   0 1 2 3 4 5 6 7 8 9 10

6. What grade did you think you would achieve on the test…

   a. Before you wrote it?

   b. While writing it?

   c. After writing it?

OBJECTIVE DEMANDS OF THE TEST AND ABILITY TO COPE

Thank you for visualizing that. I am going to ask some more questions about your experience, but you no longer need to answer as if you are in back in that point in time.

<u>Objective Demands</u>

[Note: Most of the objective demands of the test would have been collected in the "Background Information About Recent Test Anxiety Experience" section].

1. Were you trying to achieve honour roll or win a scholarship or an award for which your grades were being considered while you were feeling anxious about this test?

2. Were you in the process of applying for co-op, a job, or another program (e.g., graduate school) for which your grades were being considered while you were feeling anxious about this test?

3. Did you actually experience any of the consequences that you were worried about as a result of the grade you achieved on the test? [Prompt: Remind the participant of the perceived consequences reported earlier if needed].

4. Did you have other opportunities to make up your mark in the course if you did not do well on the test?

5. Looking back now that you are no longer feeling anxious about this test, would you say that your anxiety about this particular test was excessive or out of proportion considering how challenging and important it actually was? If so, why? If not, why?

a. Now that you have written the test, how challenging was it on a scale of 0 (not challenging at all) to 10 (extremely challenging)?

0 1 2 3 4 5 6 7 8 9 10

b. Now that you have written the test, how important do you think the test was for you **overall** on a scale of 0 (not important at all) to 10 (extremely important)?

0 1 2 3 4 5 6 7 8 9 10

6. **[Examiner rating of test demands]**

0 1 2 3 4 5 6 7 8 9 10

Objective Ability to Cope

1. Walk me through how you prepared or studied for the test? Tell me about all the strategies that you can remember using.

    a. How long before the test did you learn about it?

    b. How long before the test did you begin studying for it?

    c. Approximately how much time did you spend preparing for the test in total?

        i. How does this compare to the time you typically spend preparing for other tests?

2. When you were feeling anxious while studying for the test, how did you cope (e.g., continue studying, procrastinate)? Tell me about as many coping strategies as you can remember using.

3. Walk me through how you wrote the test. Tell me about all the strategies that you can remember using.

    a. How long did you take to write the test?

4. When you were feeling anxious while writing the test, how did you cope? Tell me about as many coping strategies as you can remember using.

5. When you are not feeling anxious about a particular test, how confident are you in your ability to effectively study and write tests **overall** on a scale of 0 (not confident at all) to 10 (extremely confident)?

0 1 2 3 4 5 6 7 8 9 10

6. **[Examiner rating of coping abilities]**

0 1 2 3 4 5 6 7 8 9 10

IMPACT AND SEVERITY OF TEST ANXIETY

I am now going to ask you some questions about different ways your anxiety about this test might have impacted you.

Academic Impact

1. Did anxiety about this test impact your…

	✓/✗	Description of impact
Grades?		
Ability to attend class?		
Ability to study for the test?		
Ability to write the test?		
Ability to complete the test on time?		
Performance on the test?		
Ability to complete other coursework?		
Performance on other coursework or in other courses?		

Social Impact

2. Did anxiety about this test impact your…

	✓/✗	Description of impact
Ability to make friends in the course?		
Ability to study with friends/peers in the course?		
Relationship with friends/peers inside/outside of the course?		
Relationship with family members?		

Relationship with other important people in your life?		
Ability to speak to the course instructor or teaching assistant(s)?		

Impact on Well-Being

3. Did anxiety about this test impact your…

	✓/✗	Description of impact
Sleep?		
Eating?		
Peace of mind?		
Irritability?		
Concentration?		

Other Impacts

4. Did anxiety about this test impact you in any other ways?

General Impact

5. Has test anxiety **in general** impacted your…

	✓/✗	Description of impact
Choice of major/minor?		
Choice of courses?		
Ability to achieve honour roll or to win a scholarship or an award?		
Ability to obtain a co-op, job, or be accepted into another program?		
Choice of career path?		
Relationships with peers, friends, family members, or other important people in your life?		

6. How would you rate your test anxiety **overall** in terms of how much it negatively impacts your life (academically, socially, etc.) on a scale of 0 (not impactful at all) to 10 (extremely impactful)?

   0 1 2 3 4 5 6 7 8 9 10

   a. Why did you rate this as a(n) _____?

7. How would you rate your test anxiety **overall** in terms of how distressing it is to you (i.e., how much does it bother you when you are feeling anxious about a test) on a scale of 0 (not distressing at all) to 10 (extremely distressing)?

   0 1 2 3 4 5 6 7 8 9 10

   a. Why did you rate this as a(n) _____?

8. **[Examiner rating of test anxiety overall]**

   0 1 2 3 4 5 6 7 8 9 10

HISTORY OF TEST ANXIETY

The last set of questions I have for you will ask about your experience with test anxiety more generally.

1. Do you experience test anxiety often?

   a. Is the experience you described earlier typical for you when tests come up? If not, how is the experience you described different from your typical experiences with tests?

   b. How many times in any given term would you say that you experience test anxiety that is similar to the experience you described earlier (e.g., once, twice, all the time)?

2. When did you first begin to experience test anxiety (e.g., elementary school, secondary school, post-secondary school)?

3. Since you first began experiencing test anxiety has it changed at all (e.g., become better/worse or more intense/less intense)?

4. Since this began, have you ever sought help for it? If so, what kinds of help have you received?

   a. Have you sought any academic or other accommodations for test anxiety in post-secondary school? What accommodations have you received, if any?

5. Do you tend to experience difficulties with anxiety in general (e.g., being really self-conscious in front of others, worrying a lot about a wide range of things, experiencing panic attacks and worrying about them)?

    a. What problems does this anxiety cause for you?

6. Have you sought any kind of treatment for anxiety in general? If so, what treatment(s)?

    a. Have you sought any academic or other accommodations for anxiety in general in post-secondary school? What accommodations have you received, if any?

CONCLUSION

1. Is there anything about your test anxiety that we have not talked about?

2. What has the experience of talking about your test anxiety been like for you?

That is the last of my questions. Test anxiety can be very difficult to deal with and talk about, so I want to thank you for your openness in sharing your experience today.