# A longitudinal study of test-related emotions in school aged secondary education students in England.

A thesis submitted to the University of Manchester for the degree of Doctorate in Educational and Child Psychology in the Faculty of Humanities 2023

**Anthony Collins**

Word count: 21021 (excluding references and tables)

School of Environment, Education and Development (SEED)

**List of Contents**

**Paper 2: Testing Times – A Longitudinal Study of Test Related Emotions**

**Paper 3: The Dissemination of Research Evidence to Professional Practice**

**Part 1**



**Part 2**

**List of Tables**

**Paper 1**



**Paper 2**



**Paper 3**

## List of Figures

### Paper 1



### Paper 2



## List of Appendices

## Thesis Abstract

Anxiety about examinations is a recognised and widespread problem, with serious implications for school-aged children. Government bodies have called for reliable measures and a better understanding of test anxiety to gather clear data, drive change, and improve practices in schools.

This research provides a contribution to knowledge and understanding of test anxiety and wider related emotions. In Paper 1, the PRISMA framework is used to systematically review the literature on test anxiety and identify the measures used in studies published from 2000 to 2020. The findings reveal issues with the reliability of current evidence due to outdated test anxiety measures and their supporting standardisation groups. There are also concerns about the validity of the measures due to the wide variation in the language used to describe different aspects of test anxiety. Additionally, the research highlights a lack of incorporation of young people's views and perspectives on test anxiety.

Paper 2 presents the findings of a quantitative, longitudinal study conducted in two secondary schools located in the Northwest of England. The study examines the levels of test emotions and anxiety in 15-16-year-old students throughout their final year of secondary education. The findings show that while some test-related emotions remain stable over time, others, such as test anxiety, test relief, test shame, and test hopelessness, fluctuate in response to significant testing events like mock and final exams. The study also reveals that young people in different school settings experience test emotions differently, possibly due to variations in school culture and broader contextual factors.

Paper 3 reflects on the implications and practical value of the research for educational psychologists and schools. It also discusses the dissemination plans to reach psychologists, teachers, and students.

Keywords: test emotions, test anxiety, educational assessment, high school, secondary school, students, well-being.

**Declaration**

I declare that no portion of the work referred to in this thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

**Funding Body Statement**

*This project was funded through England's Department for Education (DfE) ITEP award 2020-2021.*

**Copyright statement**

*(1) The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the "Copyright") and he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.*

*(2) Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made only in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.*

*(3) The ownership of certain Copyright, patents, designs, trademarks, and other intellectual property (the "Intellectual Property") and any reproductions of copyright works in the thesis, for example graphs and tables ("Reproductions"), which may be described in this thesis, may not be owned by the author, and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.*

*(4). Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see http://documents.manchester.ac.uk/DocuInfo.aspx?DocID=24420), in any relevant Thesis restriction declarations deposited in the University Library, The University Library's regulations (see http://www.library.manchester.ac.uk/about/regulations/) and in The University's policy on Presentation of Theses.*

**Acknowledgements**

I must express my immense gratitude to both Professor Kevin Woods and Doctor Tee Mcaldin, whose remarkable patience, unwavering support, impeccable guidance, and extraordinary knack for discovering both marvel and mirth throughout the entirety of the research journey have been nothing short of astonishing.

I also extend my heartfelt thanks to my partner, Ian, whose assistance, support, tolerance, and unconditional love have been my pillars of strength throughout both the process of authoring this thesis, as well as navigating the wider challenges of the Doctorate programme.

To Ann and Mary Ellis, to whom I dedicate this thesis, although unable to witness the culmination of this work, their tireless years of dedication, love, and nurture has played an indispensable role in bringing me to this juncture. You still lead me on.

## Thesis Introduction

Testing and assessments play a crucial role in educational systems as a means to evaluate students' knowledge and abilities across different subjects. While some students report that a level of heightened emotion is facilitative in the testing process (Putwain, 2014), wider research indicates that testing can have detrimental effects on students' emotional well-being, impacting their performance and long-term mental health. Test anxiety can lead to difficulties in attention and cognitive processing, negative thoughts, low self-esteem, and even physical health problems. Unfortunately, research evidence indicates that educational professionals often do not identify test anxiety and exam stress as detrimental to their school's examination results (Putwain, ibid; Pekrun et al., 2018).

As defined by scholars such as Zeidner (1998) and Cizek & Burg (2006), test anxiety is a specific form of anxiety that arises from exposure to evaluative situations. It is characterised by cognitive aspects, such as worry about failure, fear of negative consequences, and concerns about peer comparison. Additionally, it encompasses subjective feelings of tension and nervousness, as well as physical or autonomic states of excitement, such as palpitations (Spielberger & Vagg, 1995).

The proportion of young people reported to experience test anxiety varies considerably, possibly due to several factors. These include a lack of routine screening, the reliability of testing measures, and the differing thresholds for reliable differentiation (Howard, 2020). The literature suggests that test anxiety could affect up to 40% of school-age populations (von der Embse, Barterian & Segool, 2013). More recent UK-based research by Putwain & Daly (2014) suggests that up to 16% of children and young people in Year 11 cohorts may suffer high TA levels. If this is the case, the most recent DfE (2019) school census data suggest that this could potentially mean that around 77 students in an average-size UK Y11 school cohort might experience test anxiety. Of that group, 32 might experience high levels of test anxiety.

Given these potential prevalence rates, Howard (2020) argues the need for further research to better understand test anxiety and its possible wide-reaching effects on children and young people's academic performance, self-confidence, and mental health. She stresses that the key to this research is the need for a reliable measure of high levels of test anxiety that can accurately assess the individual's levels and appropriately measure any response to possible interventions.

While there is a wide variety of measures available to gauge test anxiety levels, several of which date back to the 1960s, initial research indicated that there is no available review that sets out which instruments have been used within research, or more importantly examines these measures' underpinning validity or reliability. Further, given the requirement for person-centred practice within the SEND Code of Practice (DFE, 2014) and the Health & Care Professions Council (HCPC) standards of proficiency for practising educational psychologists (HCPC, 2017), it is appropriate that measures of test anxiety be re-examined to understand the level to which they match with contemporary understanding around test anxiety and the degree to which children and young people's views, experiences and judgements have been sought and incorporated into their design and interpretation.

However, measuring test anxiety at one fixed point may not prove to be particularly enlightening. The existing literature on test emotions reveals significant variations in the reported prevalence, indicating that test emotions are not consistent and static. Instead, they fluctuate over time in response to various factors, such as the increasing academic pressures students encounter throughout the academic year (Pekrun, Elliot, and Maier, 2009). It is also possible that these are influenced by individual contextual factors, such as previous academic achievement, cognitive ability, as well as personality traits (Schneider & Preckel, 2017).

Given the context outlined here, this research has two distinct aims; first, to identify which existing test anxiety measures have been used within research and review them in line with

current scale development and use recommendations. Second, to investigate the pattern of test emotions and how they vary over time, as well as to identify possible. contextual factors within the school setting that may influence test emotions.

### *The researcher's professional background*

The researcher's professional journey began after qualifying as a secondary school teacher in 1997. They accumulated over fifteen years of experience in mainstream and special educational needs settings in Northwest England, fulfilling roles such as teacher, Special Educational Needs Coordinator (SENCo), and Assistant Headteacher. In 2014, they expanded their horizons by working internationally, initially as the Head of a Secondary School in Berlin, Germany, and subsequently in Marbella, Spain.

Throughout their career, particularly as school leader for special educational needs, the researcher has first-hand experience supporting children and young people who face difficulties with test anxiety and test-related emotions. However, their observations suggested that a lack of awareness and resource limitations often prioritises other areas, such as Key Stage progress levels and GCSE pass rates, over addressing social, emotional, and mental health needs in schools. Nevertheless, it is evident that many children and young people encounter challenges with test anxiety, and for some, this pressure poses significant emotional difficulties.

The researcher's experience in schools and their understanding of the systems and dynamics within secondary school settings played a key role in navigating the research process. Their practical knowledge also aided the research process by helping them to identify and proactively address potential research challenges, ensuring a smoother execution of the study.

### *Research philosophy*

As a trainee educational psychologist, educator, and researcher, the researcher's axiological

position centres around advocating for children. It is their view that children and young people should be considered credible sources, and their voices and perspectives should be at the forefront of decision-making and policy. As such, these research endeavours are driven by the goal of supporting the well-being and empowerment of children. The researcher is mindful of their positionality and has worked to address power imbalances that may exist within the research process. This involved actively engaging with children and young people, respecting their autonomy, and adhering to ethical practices at all stages of the research journey.

Paper 1 is a literature review that examines the measures of test anxiety employed in research involving school-aged children from 2000 to 2020. The paper also raises the question of whether these identified measures are both valid and reliable, and whether they effectively capture children and young people's current perspectives, experiences, and interpretations of test anxiety.

Paper 2 expands upon the findings of Paper 1 by presenting a quantitative, longitudinal study that investigates the changes in test emotions among Year 11 students throughout the academic year. Additionally, it explores the extent to which these emotions vary across different school settings.

Both papers take a relativist ontological stance, as decribed by Cohen, Manion, and Morrison (2011). A relativist ontological position as the underpinning aim of this study as it will not be a search for a single, objective '*truth'*; instead, this paper the utility of the various measures (Cohen, Manion & Morrison, 2011; Jones & Forshaw, 2012). By taking a critically relativist epistemological position, it is accepted that while there may be a 'veracity around measuring test anxiety', all measures and their use are influenced by both experience and interpretation.

The overall endeavour of this research starts from a relativist ontological stance which intends to broaden perspectives from test anxiety as a singular construct to a range of possible test emotions and to offer a perspective for discussion about the range and variation of individuals responses and on change over time. While Paper 2 employs a quantitative, closed measure, which could be perceived as leaning towards a realist position, the interpretation of the data and the overall intent of the research demonstrate that the researcher is not attempting to promote a simplified, universally applicable truth that applies to all schools.

***References***


Cizek, G. J., & Burg, S. S. (2006). *Addressing test anxiety in a high-stakes environment: Strategies for classroom and schools.* Corwin Press.

Cohen, L., Manion, L. & Morrision, K. (2011). *Research methods in education*. London: Routledge

Department for Education and Department of Health (2015) Special educational needs and disability code of practice: 0 to 25 years. Available at: https://www.gov.uk/government/publications/send-code-of-practice-0-to-25 (Accessed: 25 March 2021).

Howard, E. (2020). *A review of the literature concerning anxiety for educational assessme*nts. Ofqual: Coventry.

Pekrun, R., Muis, K.R., Frenzel, A. C., & Goetz, T. (2018). *Emotional at School*. New York: Routledge.

Pekrun, R., Elliot, A. J., & Maier, M. A. (2009). Achievement goals and achievement emotions: Testing a model of their joint relations with academic performance. *Journal of Educational Psychology*, 101(1), 115-135.

Putwain, D.W. & Daly, A. (2014) Test anxiety prevalence and gender differences in a sample of English secondary school students. *Educational Studies*, 40:5, 554-570

Schneider, M., & Preckel, F. (2017). Variables associated with achievement in higher education: A systematic review of meta-analyses. *Psychological Bulletin*, 143(6), 565-600.

Spielberger, C. D., & Vagg, P. R. (Eds.). (1995). *Test anxiety: Theory, assessment, and treatment*. New York: Taylor & Francis.

Stobart, G. (2008), *Testing times: The uses and abuses of assessment*. Routledge, London.

Ursache, A., Raver, C.C., (2014) Trait and state anxiety: relations to executive functioning in an at-risk sample. *Cognition & Emotions.* 28(5):845-55.

Von der Embse, N., Barterian, J. and Segool, N. (2013) Test Anxiety Interventions for Children and Adolescents: A Systematic Review of Treatment Studies from 2000–2010. *Psychological Schools Journal.* 50: 57–71.

Zeidner, M. (1998). *Perspectives on individual differences: Test anxiety - The state of the art.* Plenum Press, London.

# Paper 1: Test Anxiety – A review of measures used in test anxiety research.

Paper prepared in accordance with the author guideline and formatting template for the journal, Applied Measurement in Education (See Appendix 1).

Word count: 5154 (excluding references and tables)

## Abstract

Given the potential negative implications for children and young people who experience high levels of test anxiety, Howard (2020) suggests that a reliable measure has the potential to provide a better understanding of prevalence levels and inform the development of supportive pastoral practice in schools. However, whilst there is a wide variety of measures available to gauge test anxiety levels, there is no available review that sets out which instruments have been used within research, or more importantly examines these measures' underpinning validity or reliability.

This literature review follows the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) framework (Moher, Liberati, Tetzlaff, & Altman, 2009) to systematically search for and identify the various test anxiety measures used to gauge test anxiety in the school-age population across the last twenty years. The findings of this paper indicate that all such studies used self-reporting questionnaire-type measures. Half the identified studies used test anxiety measures that used norm standards that dated from the 1970s. Nearly a quarter of research papers published in English (30 out of 127) used measures of test anxiety with school-aged children that were initially developed for use with university undergraduates. Only one measure showed evidence of children's and young people's personal views and experiences of test anxiety being sought, gathered, and incorporated as part of its developmental procedure. Furthermore, the uneven geographic spread of countries originating research papers in this area indicates 'pockets' of specialist researchers and broader cultural differences around test anxiety that may be influenced by differing national incidence levels, health priorities, and the development of, or reach within, educational systems.

It is concluded that the current evidence base around test anxiety is compromised through a weakness in the underpinning measures' convergent and content validities. Further research is needed to contemporaneously validate existing measures with children and young people's views, experiences, and understanding of test anxiety.

**Keywords:** test anxiety, measures, school, international, child-voice, validity.

**Introduction.**

Test anxiety (TA) is considered a specific form of anxiety resulting from exposure to testing or similar evaluative situations (Zeidner, 1998; Cizek & Burg, 2006). Current understanding proposes that it is characterised by: (a) a worry component which refers to the cognitive aspects experienced in evaluative situations such as worry about failure and fears around possible implications and negative peer comparison, (b) an emotionality aspect which relates to the subjective feelings of tension (e.g., nervousness) and, (c) physical or autonomic states of excitement (e.g., palpitations) (Spielberger & Vagg, 1995).

There is a considerable body of research around test anxiety dating back to the 1950s. This body of literature discusses the development of theories around possible underpinning conceptualisations, its impact on academic performance as well as possible interventions that may support young people through the testing cycle. Early models of TA explained the worries and fears children experienced around school grading and examinations either as being the result of phobias and neuroses or as being linked to levels of general anxiety (Mander & Sarason, 1952; Lapouse & Monk, 1959; Eysenck & Rachman, 1965). Subsequent research refined understanding away from these unidimensional perspectives by differentiating test anxiety from general anxiety and defining it as a situationally specific trait whereby individuals are predisposed to experience anxiety in advance of, during and following performance-evaluative situations (Zeidner, 1998; Spielberger & Vagg, 1995).

Within current literature, there is a consensus around test anxiety being a distinct form of anxiety, although there remains a lack of agreement, conceptual clarity and uncertainty regarding just how distinct test anxiety is as a characteristic (Cizek & Burg, 2006). Indeed, more recent models of TA point towards a complex, multi-dimensional construct that involves an ever-increasing range of factors that encompass poor self-perception of academic competence, aspects of social humiliation, as well as elements of social derogation (Zeidner

& Matthews, 2005; Segool et al., 2014). Putwain et al. (2020) highlight this lack of consensus around the dimensionalities of test anxiety, arguing that several of these factors may simply be different terms that describe very specific elements of the wider dimensions of worry and emotionality.

**The impact of test anxiety on children and young people.**

Putting aside the debate around the underpinning construct of test anxiety, it is known that although some students find a level of heightened emotion to be facilitative, higher levels of anxiety in children are known to be associated with attentional and cognitive processing difficulties that can have a detrimental impact on academic performance (Hembree, 1988; Seipp, 1991; Cassady & Johnson, 2002; Keogh et al., 2004; Ursache & Raver, 2014; Ajilchi & Nejati, 2017). In addition, anxiety levels that focus on testing and assessment appear to negatively influence self-esteem (Peleg, 2009; Thomas & Gadbois, 2007). Other studies have shown that long-term exposure to higher anxiety levels is linked to disrupted sleep patterns, higher risk of mental health difficulties, and increased exposure to various illnesses resulting from suppression of the immune system (Damer & Melendres, 2011; Mental Health Foundation, 2014). Perhaps the most concerning findings are that in the adolescent population, high levels of examination pressures have been implicated as a significant antecedent factor in incidents of self-harm and suicide (Rodway et al., 2016). Accordingly, given the negative impact that test anxiety may have on children and young people, this has led educationalists and professionals to explore ways to mitigate its effects, leading to the development of a range of interventions such as STEPS – Strategies to Tackle Exam Pressure and Stress (Putwain, Chamberlain, Daly, and Sadreddini, 2014) and BEAT – Beating Exam Anxiety Together (Kent Educational Psychology Service, 2016).

**Prevalence of test anxiety.**

The proportion of young people reported to experience test anxiety varies considerably,

possibly due to several factors. These include primarily a lack of routine screening, the reliability of testing measures, and the differing thresholds for reliable differentiation (Howard, 2020). Significantly, the literature suggests that test anxiety could affect up to 40% of school-age populations (von der Embse, Barterian & Segool, 2013). More recent UK-based research by Putwain & Daly (2014) suggests that up to 16% of children and young people in Year 11 cohorts may suffer high TA levels. If this is the case, the most recent DfE (2019) school census data suggest that this could potentially mean that around 77 students in an average-size UK Y11 school cohort might experience test anxiety. Of that group, 32 might experience high levels of test anxiety. However, test anxiety remains largely unrecognised in schools in the UK and internationally (Putwain, 2014; Pekrun, 2018).

**The need for effective measures of test anxiety.**

Given these potential prevalence rates, Howard (2020) argues the need for further research to better understand test anxiety and its possible wide-reaching effects on children and young people's academic performance, self-confidence, and mental health. She stresses that the key to this research is the need for a reliable measure of high levels of test anxiety that can accurately assess the individual's levels and appropriately measure any response to interventions intended to reduce test anxiety.

However, whilst there is a wide variety of measures available to gauge test anxiety levels, several of which date back to the 1960s, at the time of writing, there is no available review that sets out which instruments have been used within research, or more importantly examines these measures' underpinning validity or reliability. Further, given the requirement for person-centred practice within the SEND Code of Practice (DFE, 2014) and the Health & Care Professions Council (HCPC) standards of proficiency for practising educational psychologists (HCPC, 2017), it is appropriate that measures of test anxiety be re-examined to understand the level to which they match with contemporary understanding around test

anxiety and the degree to which children and young people's views, experiences and judgements have been sought and incorporated into their design and interpretation.

Indeed, this links closely to a wider body of research that points to concerns around the development and use of instruments that aim to measure elements of children and young people's phenomenologically experienced difficulties not always considering or including the views and experiences of young people (Butler & Gasson, 2005; Byrne, 2002; Harter, 1999; Wylie, 1989). Tatlow-Golden & Guerin (2015) highlight that children and young people's contemporary views, experiences, and meanings around self-concepts such as test anxiety are rarely explored and included in design and construction psychological self-concept measures. As such they suggest that concepts, such as test anxiety, are likely to consist of multiple salient factors which are often not included within measures and, in doing so, these omissions raise important questions about the content validity and uses of such measures.

**Research aims.**

Given the need for a reliable measure of test anxiety, the lack of a review of measures as well as concerns around validity, applicability and reliability, this research sets out to identify which existing test anxiety measures have been used within research and reviews them in line with current scale development and use recommendations. Specifically, the research questions are:

(1) What measures of test anxiety have been used to inform research in the period 2000-2020?
(2) Do the measures used to inform test anxiety research provide valid and reliable evidence of children and young people's contemporary views, experiences, and meanings around test anxiety.

## Method

### *Ontological & Epistemological Position.*

The research questions above are designed to explore, examine, and explain the various measures of test anxiety used in contemporary research. As such, it is appropriate to take a relativist ontological position as the underpinning aim of this study as it will not be a search for a single, objective '*truth'*; instead, this paper the utility of the various measures (Cohen, Manion & Morrison, 2011; Jones & Forshaw, 2012). By taking a critically relativist epistemological position, it is accepted that while there may be a 'veracity around measuring test anxiety', all measures and their use are influenced by both experience and interpretation.

### *Study Design.*

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) framework (Moher, Liberati, Tetzlaff, & Altman, 2009) was used to identify and select appropriate papers for review. This approach was selected following a discussion between the researcher and research supervisor to ensure a rigorous and reproducible study. Between August 2021 and September 2022, the following databases were systematically searched for relevant studies: Psych Info, Education Resources Information Center (ERIC), Medinc, and the University of Manchester Online library.

Although the terms 'test anxiety' and 'examination stress' are sometimes used interchangeably within professional practice, the term 'examination stress' was excluded from this study as the researcher judged it to be a limiting term that related to specifically to examination situations. In contrast, 'test anxiety' encompasses multiple, broader, and contextual features within a child-focussed perspective. This approach is supported by Mccaldin's (2019) findings, where examination stress was greatest prior to the commencement of the annual exam season and decreased rapidly once it had begun. Consequently, key search terms included 'test anxiety', 'exam anxiety', 'test anxiety

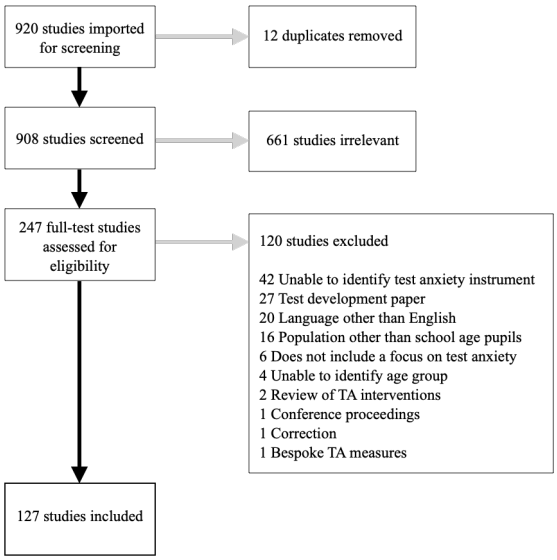
measures', and 'test anxiety instruments'. Searches were conducted using both single and combined terms. Additionally, reference harvesting from relevant articles was also carried out.

The search generated 920 papers, of which 661 were excluded after removing duplicates and screening titles and abstracts. The remaining 247 potentially relevant papers were screened against the following inclusion criteria:

(3) Published between January 2000 and December 2021
(4) Written in English
(5) Focus on test anxiety measurement.
(6) Study is conducted with school-age students (4–18 years old)
(7) Subjected to peer review in an academic journal.
(8) Was not related to the development of a test anxiety measure.
(9) Was not a conference proceeding.

At this point, a further 120 papers were excluded as they did not fully address the inclusion criteria. Details of the excluded papers can be found in Figure 1 below which shows the PRISMA diagram for the study. The research supervisor then screened a random sample of included studies to ensure that the reliability of the sample.

***Figure 1: PRISMA diagram.***




***Data extraction, analysis, and synthesis.***

The researcher read each paper, extracted the core features from each study, and compiled them into Table 1 below), outlining the measures used for the various studies, the authors of the measures, and the number of studies in which the measure was used. Further research was required to gather the broader details of the various test anxiety measures identified, including their developmental purpose and claims to validity and reliability. To do this in a consistent, repeatable, and structured way, the researcher developed an analytic framework to examine the test anxiety measures. This framework is presented in Appendix 2 and captures the available information on a range of areas, including details of the standardisation sample, administration time, reliability, and validity information. Following an initial draft, the framework was verified and refined following a discussion between the researcher and the research supervisor.

The data gathered from through the analytic framework were compiled to generate Table 1. This sets out a comprehensive list of all the measures, their underpinning constructs, sample group norms, and validity and reliability details. As part of the analysis and synthesis process, the researcher examined the published manuals or papers describing the construction of the various measures to evaluate the construct validity across all the measures. The collective

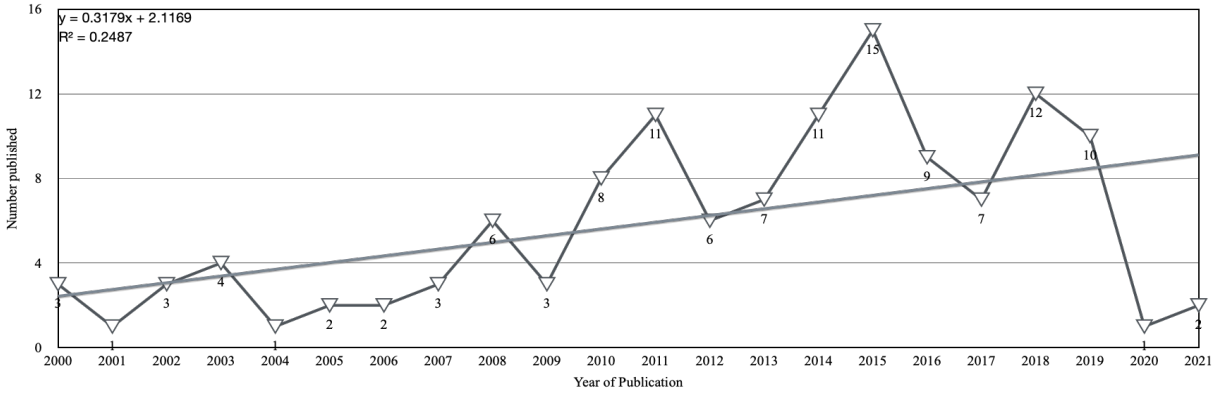
terms were then grouped into meaningful domains presented within Table 2. Again, following this initial draft, the framework was verified and refined following a discussion between the researcher and research supervisor.

**Findings.**

***Research question 1: What measures of test anxiety have been used to inform research in the period 2000- 2020?***

In total, 127 papers, published in English, were found to have examined test anxiety within the school-age population between January 2000 and December 2020, with the mean number of studies per year being 5.9 and the range of studies per year was between 1 and 16. In the first decade, from 2000 to 2010, 36 studies were identified, which was found to have increased to 98 in the following decade. Figure 2 below indicates that the number of papers published in 2020 and 2021 dropped significantly below the mean. A linear trend line indicates an increasing interest in publishing test anxiety papers in English in the school-age population.

***Figure 2: Annual number of identified school aged test anxiety papers (Sept 2000 - Dec 2021)***




Twenty-seven studies (21%) were found to have been conducted with primary-aged students (Ages 5-11), 99 (78%) were conducted with secondary-aged students (See table below 1)

(Ages 11-18), and one study (0.8%) had been conducted across both educational phases (Ages 5-18).

Table 1 lists the 20, diverse test anxiety measures identified as having been used with school-age students across the time period. Two measures, the Test Anxiety Inventory (34.65%) and the Children's Test Anxiety Scale (14.17%), were found to have been used within nearly 50 per cent of the studies. Table 1 also includes details of available information concerning the validity and reliability of the measures. In general, it was found that there was regular evidence of the measure's validity and reliability in conventional terms, though only one measure was found to include details of how children and young people were included in the development of the measure.

A full breakdown of the studies by country is provided in Figure 3 below. However, analysis indicates three countries leading test anxiety research since 2000. These are the USA (31 studies), the UK (20 studies) and Germany (15 studies). Other countries have also produced higher than average numbers of studies, e.g., Singapore and Israel. It is not known why there is an uneven geographic spread of countries originating research papers in this area. However, it provides an indication of *'pockets'* of specialist researchers as well as the possibility of wider cultural differences around test anxiety that may be influenced by factors such as differing national incidence levels, health priorities, and the development of, or reach within educational systems.

***Figure 3: Chart Showing the Number of Studies Published in English by Country (2000 - 2021)***

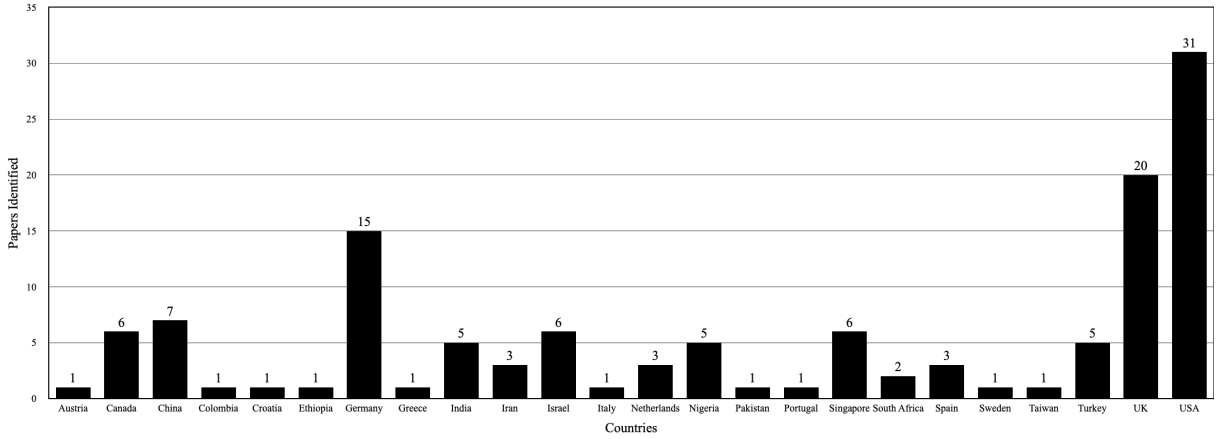


***Table 1: Test anxiety measures used in research between January 2000 and December 2021.***

| Test Anxiety Measure/ Authors/ Date of norms | Age range/ Approx admin time | Number/ Percentage of Studies/ Countries | Purpose/ Underlying constructs focus. | Description | Validity | Reliability | Underpinning construct validated with norm group | Scale outputs |
|---|---|---|---|---|---|---|---|---|
| Achievement Motivation Questionnaire for Students<br><br>Franz Petermann & Sandra Winkel<br><br>2017 | 12 – 19<br><br>30 mins | 3 (2.36%)<br><br>Germany | Aims to explore feelings relating to both inhibiting/ debilitating and facilitative effects.<br><br>Worry, fear, motivation, effort. | Self-report Likert questionnaire; 30 items; to be completed before and during evaluation situations. | Correlations to AMS-R, AMG-S and SELLMO highly significant. CFA: Five scales correspond to five factors accounting for 98% of variance. | Internal consistency: r = .62 and r = .74. Test-retest reliability: r = .67 and r = .76. | No data provided | T-Scores and percentiles for males and females, years 7 (11yrs) to 13 (18yrs) for each of five scales. |
| Anxiety Questionnaire for Students<br><br>W. Wieczerkowski<br>H. Nickel<br>A. Janowski<br>B. Fittkau<br>W. Rauer<br><br>1975 | 9 -19<br><br>50 mins | 1 (0.78%)<br><br>Germany | The Anxiety Questionnaire for Students was developed as a measure for test anxiety in German schools.<br><br>Test anxiety, manifest anxiety, dislike of school and social desirability. | Self-report Likert questionnaire; 50 items; to be completed before evaluation situations. | Correlation with the Test Anxiety Inventory (TAI). Validity coefficient is given as r = 0.80. | No reliability data is provided. | No data provided | Cut off points for descriptive ranges are provided based on mean scores. |
| Bakare Test Anxiety Scale<br><br>Christopher G Bakare<br><br>1969 | 18+<br><br>40 mins | 1 (0.78%)<br><br>Nigeria | Developed as a Nigerian measure of test anxiety.<br><br>Worry, Physiological Symptoms. | Self-report Likert questionnaire; 37 items; to be completed before evaluation situations. | No data available | Test-retest reliability: r = 0.89. | No data provided | Cut off points for descriptive ranges are provided based on summed scores. |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Children's Test Anxiety Scale<br><br>Douglas G. Wren<br>Jeri Benson<br><br>2002 | 8 – 12<br><br>30 mins | 18 (14.17%)<br><br>China<br>Netherlands<br>Singapore<br>Spain<br>Sweden<br>UK<br>USA | Aims to provide a reliable self-report instrument for use with US children that includes some ethnic groups with the sample (Cuban, Mexican)<br><br>Negative thoughts, off task behaviour and autonomic reactions. | Self-report Likert questionnaire; 30 items; to be completed before evaluation situations. | Based upon an item analysis, coefficient alpha for the reduced 30-item instrument was 0.92, and the subscale alphas ranged from 0.78 to 0.89. | The reliability estimates for the 30-item CTAS was 0.92 and its subscales were 0.85 for Autonomic Reactions, 0.78 for Off-Task Behaviours, and 0.89 for the Thoughts subscale. | No data provided | T-Scores and percentiles are available for males and females for US Grades 3 (8yrs) to 6 (12yrs) for the five scales. |
| Cognitive Test Anxiety Scale – 2nd Edition.<br><br>Christopher L. Thomas<br>Jerrell C. Cassady<br>Wendy Holmes Finch<br><br>2016 | 18+<br><br>25 mins | 1 (0.78%)<br><br>Greece | Aims to explore the cognitive indicators of test anxiety.<br><br>Worry and emotionality. | Self-report Likert questionnaire; 24 items; to be completed before and during evaluation situations. | Correlations to the Motivated Strategies for Learning Questionnaire (MSLQ) and and Friedben Test Anxiety Scale. Validity coefficient is given as r = 0.81. | No reliability data is provided. | No data provided | Cut off points for descriptive ranges are provided based on mean scores are provided for both males and females. |
| Examination Stress Scale<br><br>Yao-Ting Sung<br>Tzu-Yang Chao<br><br>2015 | 13 – 16<br><br>25 mins | 1 (0.78%)<br><br>Taiwan | Aims to explore the examination anxiety through a stress framework.<br><br>Physiological Anxiety Responses, Cognitive and behavioural Responses, Perceived Social Expectation and Social Comparison. | Self-report Likert questionnaire; 27 items; to be completed before and during evaluation situations. | Correlation with the Test Anxiety Inventory (TAI). Validity coefficient is given as r = 0.77. | Internal consistency: r = .89, r = .85 and r = .88. Test-retest reliability not reported. | No data provided | Cut off points for descriptive ranges are provided based on summed scores. |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Friedben Test Anxiety Scale<br><br>Isaac A Friedman<br>Orit Bendas-Jacob<br><br>1997 | 11 -16<br><br>25 mins | 4 (3.15%)<br><br>Israel<br>USA | Aims to be a broad measure of test anxiety for use with adolescents.<br><br>General Fear and Worrying Thoughts, Tension and Restlessness (Emotionality), Threat to Perceived Self-Image or Self-Efficacy due to Failure, Threat to Social Status. | Self-report Likert questionnaire; 23 items; to be completed before evaluation situations | Correlations to the Test Anxiety Inventory (TAI). Validity coefficient is given as r = .81, and the subscale alphas ranged from 0.81 to 0.91. | Internal consistency: r = .78 and r = .83. Test-retest reliability: r = .54 and r = .76. | Authors report that the underpinning constructs were developed from and checked with 80 Grade 11 students (age 17). | Cut off points for descriptive ranges are provided based on summed scores |
| Multidimensional Test Anxiety Scale<br><br>David Putwain, Nathaniel von der Embse, Emma Rainbird, Geoffrey West<br><br>2012 | 11 – 16<br><br>15 mins | 1 (0.78%)<br><br>UK | Aims to reflect the multidimensionality of test anxiety by including four separate scales that that the authors describe as making up test anxiety.<br><br>Worry, cognitive interference, tension, and physiological indicators. | Self-report Likert questionnaire; 15 items; to be completed before evaluation situations. | Correlations to Test Anxiety Inventory (TAI), School Related Well-Being Scale (SRWS) and the Social, Academic, and Emotional, Behaviour Risk Screener – Student Risk Scale (SAEBRS-SRS). | Test-retest reliability; r = .80 for the MTAS total score, r = .80 for Worry, r = .65 for Cognitive Interference, r = .70 for Tension, r = .82 for Physiological Indicators. | No data provided | Cut off points for descriptive ranges are provided based on score distributions. |
| Revised Test Anxiety Questionnaire<br><br>Jeri Benson<br>Nabil El-Zahhar<br><br>1992 | 18+<br><br>25 mins | 2 (2.54%)<br><br>UK | Developed as a measure of test anxiety for research purposes.<br><br>Tension, worry, bodily symptoms and test irrelevant thinking. | Self-report Likert questionnaire; 25 items; to be completed before evaluation situations. | Correlations to the Test Anxiety Questionnaire (TAS). Validity coefficient is given as r = .44, and the subscale alphas ranged from 0.78 to 0.89. | No reliability data is provided. | No data provided | Means and standard deviations are provided for total scores as well as for the four subscales. |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Revised Test Anxiety Scale<br><br>J. Benson,<br>M. Moulin-Julian,<br>C. Schwarzer<br>B. Seipp, B<br>N.E. El-Zahhar<br><br>1992 | 18 - 24<br><br>18 mins | 9 (7.08%)<br><br>Israel<br>Turkey<br>UK | Developed as a multidimensional measure of test anxiety.<br><br>Tension, Worry, Bodily Symptoms, and Test-Irrelevant Thinking. | Self-report Likert questionnaire; 18 items; to be completed before evaluation situations. | Correlated with the Reactions to Tests scale (RTT). Validity coefficient is given as r = .88, and the subscale alphas ranged from 0.68 to 0.82. | No reliability data is provided. | No data provided | Cut off points for descriptive ranges are provided based on summed scores |
| State-Trait Anxiety Inventory<br><br>C.D. Speilberger<br>R.L. Gorsuch<br>R. Lushene<br>P.R. Vagg<br>G.A. Jacobs<br><br>1970 | 14+<br><br>40 mins | 6 (4.72%)<br><br>Canada<br>Germany<br>Iran | The State-Trait Anxiety Inventory is a measure of trait and state anxiety for use in clinical settings to diagnose anxiety and to distinguish it from depressive syndromes.<br><br>State anxiety, trait anxiety. | Self-report Likert questionnaire; 40 items; to be completed before evaluation situations. | The authors report internal consistency coefficients for the scale have ranged from .86 to .95. | Trait-anxiety scale test-retest reliability: r = .65 and r = .86. State-anxiety scale test-retest reliability: r = .16 and r = .62. | No data provided | Cut off points for descriptive ranges are provided based on mean scores. |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Test Anxiety Inventory<br><br>Charles D. Spielberger<br><br>1977 | 14 – 18<br><br>15 mins | 44 (34.65%)<br><br>Canada<br>China<br>Germany<br>India<br>Nigeria<br>Singapore<br>South Africa<br>Spain<br>Turkey<br>UK<br>USA | Aimed as a measure of outcomes in studies of test anxiety treatment.<br><br>Worry and emotionality. | Self-report Likert questionnaire; 20 items; to be completed before evaluation situations. | Correlation to Anxiety Scales for Children (ASC) and State-Trait Anxiety Scale (STAS) | Internal consistency: 0.92 for TAI-T, 0.91 for TAI-W and 0.91 for TAI-E. Test-retest reliability: r=.81and r=.62. | No data provided | Percentile ranks and normalised T-scores (Mean=50; SD=10) for the overall test anxiety scales and the two subscales are provided for both males and females. |
| Test Anxiety Measure for Adolescents<br><br>Patricia A. Lowe<br><br>2014 | 11 – 19<br><br>50 mins | 1 (0.78%)<br><br>USA | Developed as a multidimensional measure of test anxiety.<br><br>Behavioural, cognitive interference, physiological hyperarousal, social derogation, and worry. | Self-report Likert questionnaire; 50 items; to be completed before evaluation situations. | Correlation with the Revised Children's Manifest Anxiety Scale (RCMAS-2). Validity coefficient is given as r = 0.81. | Internal consistency: r = .87. Test-retest reliability: r = .69 and r = .78. | No data provided | Means and standard deviations are provided for total scores as well as for the five subscales. |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Test Anxiety Questionnaire<br><br>G. Mandler<br>J.E. Cowen<br><br>2001 | 14 – 18<br><br>20 mins | 18 (14.17%)<br><br>Croatia<br>Germany<br>Israel<br>Netherlands<br>Pakistan<br>UK<br>USA | Developed as a tool for use in counselling, therapy, school, and pedagogical contexts.<br><br>Excitement, anxiety, interference, and lack of confidence. | Self-report Likert questionnaire; 20 items; to be completed before evaluation situations. | Correlated to the earlier TAQs versions. Data on the factor analysis for the four areas of excitement, anxiety, interference, and confidence. Based upon an item analysis, coefficient alpha for the reduced 30-item instrument was 0.92, and the subscale alphas ranged from 0.61 to 0.70. | Reported internal consistency (total scale: .88; excitement: .86; concern: .81; interference: .79; lack of confidence: .85) and retest reliability (total scale: .86; excitement: .80; worry: .79; interference: .79; lack of confidence: .83 | No data provided | T-Scores and percentiles are available for males and females from 14 to 18 years. |
| Test Anxiety Questionnaire for Children<br><br>Marci Donolato Mammarella Altoè<br><br>2017 | 6 – 13<br><br>30 mins | 1 (0.78%)<br><br>Italy | Developed as a multidimensional measure of test anxiety for Italian children.<br><br>Anxiety, Poor Self-Evaluation, Somatic Symptoms, General School Concerns. | Self-report Likert questionnaire; 30 items; to be completed before evaluation situations. | No validity data is provided. | No reliability data is provided. | No data provided | Cut off points for descriptive ranges are provided based on mean scores. |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Test Anxiety Scale<br><br>Irwin Sarason<br><br>1976 | 18+<br><br>35 mins | 15 (11.81%)<br><br>China<br>Colombia<br>Ethiopia<br>Iran<br>Nigeria<br>South Africa<br>Turkey<br>UK<br>USA | Developed as a research instrument to identify extreme scores as well as to provide a reliable measure for further experimentation around test anxiety.<br><br>Anxiety and worry. | Self-report Likert questionnaire; 37 items; to be completed before evaluation situations. | Correlation with the Test Anxiety Inventory (TAI). Validity coefficient is given as r = 0.82. | Internal consistency: r = .87. Test-retest reliability: r = .67 and r = .76. | No data provided | Cut off points for descriptive ranges are provided based on mean scores are provided for both males and females. |
| Test Anxiety Scale for Children<br><br>Seymour B. Sarason<br>Kenneth Davidson<br>Frederick Lighthall<br>Richard Waite<br><br>1960 | 11 - 13<br><br>30 mins | 6 (4.72%)<br><br>Spain<br>USA | Aims to explore test anxiety in elementary students in USA.<br><br>Test Anxiety, Somatic Signs of Anxiety, Recitation Anxiety, Manifest Dream Anxiety. | Self-report Likert questionnaire; 30 items; to be completed before evaluation situations | The authors report a validity coefficient of .82 between the TASC and an unnamed questionnaire measure of school anxiety. | A split-half reliability is reported as being .88 and an alpha coefficient of .88 for the validation sample. | No data provided | Means and standard deviations are provided for total scores as well as for the four subscales. |
| Test Anxiety Scale for Elementary Students<br><br>Patricia A. Lowe<br>Mattew J. Grumbein,<br>Jennifer M. Raad<br><br>2008 | 7 – 12<br><br>30 mins | 1 (0.78%)<br><br>Singapore | Aims to explore the multidimensionality of test anxiety in elementary students in USA.<br><br>Physiological hyperarousal, social concerns, task irrelevant behaviour. | Self-report Likert questionnaire; 30 items; to be completed before evaluation situations | Correlation with the Test Anxiety Inventory (TAI). Validity coefficient is given as r = 0.78. | Reported internal consistency: Total scale: 0.78. Test-retest reliability not reported. | No data provided | Means and standard deviations are provided for total scores as well as for the four subscales. |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Test Anxiety Inventory for Children & Adolescents<br><br>Patricia A. Lowe Steven W. Lee Kristin M. Witteborg Keri W. Prichard Megan E. Luhr Christopher M. Cullinan Bethany A. Mildren Jennifer M. Raad Rebecca A. Cornelius Melissa Janik<br><br>2008 | 9 - 18<br><br>45 mins | 2 (1.57%)<br><br>Iran USA | Aims to explore the multidimensionality of test anxiety USA Grade 4 -12 students.<br><br>Cognitive obstruction/inattention, Physiological hyperarousal, social humiliation and lies. | Self-report Likert questionnaire; 45 items; to be completed before evaluation situations. | Correlations with the Behaviour Assessment System for Children – Self-Report of Personality, Revised Children's Manifest Anxiety Scale and the Reynolds Intellectual Screening Test. Validity coefficient is given as r = 0.80. | Internal consistency: r = .87. Test-retest reliability: r = .81 and r = .90. | No data provided | Cut off points for descriptive ranges are provided based on mean scores. |
| Westside Test Anxiety Scale<br><br>Richard Driscoll<br><br>2004 | 10 -24<br><br>10 mins | 1 (0.78%)<br><br>Turkey | Designed as a brief instrument for use by school counsellors. It aims to identify students who could benefit from anxiety- reduction interventions.<br><br>Cognitive impairment, worry and dread. | Self-report Likert questionnaire; 10 items; to be completed before evaluation situations. | Correlations to the Cognitive Test Anxiety Scale by Cassady and Johnson (2001). Validity coefficient is given as r = .44. | No reliability data is provided. | No data provided | Cut off points for descriptive ranges are provided based on mean scores. |

***Research question 2: Do the measures used to inform test anxiety research provide valid and reliable evidence of children and young people's contemporary views, experiences, and meanings around test anxiety.***

All the identified measures of test anxiety were developed using representative sample groups of varying sizes and age profiles. Of the 20 measures that were identified as being used across the time period, six were found to have been specifically developed with norm groups made up of school-aged children, the remaining 14 being developed with undergraduate sample populations. Other than the gender split between the various groups, none of the measures reported details of the samples' compositions in relation to equality or diversity profiles or participants with any identified special educational needs. Other than describing age and gender identification, only two of the measures reported broader details around the norm sample group composition concerning the social-cultural makeup of the sample groups. None of the measures referenced the inclusion of individuals with special educational needs. Of the 20 measures, 14 were found to have been developed with undergraduate students in university settings.

Only the FriedBan Test Anxiety Scale (Friedman & Bendas-Jacob, 1997) was found to reference any checking of the measure's underpinning the constructs with the norm group. In this case, Friedman and Bendas-Jacob (ibid) report that they developed the measures underpinning constructs from 80 written responses from children and young people in Grade 11 (age 17) who were asked four open-ended questions about feelings and behaviours related to test anxiety which was used to identify the underpinning constructs of social derogation, cognitive obstruction, and tenseness. Following the draft of their measure, Friedman and Bendas-Jacob (ibid) asked 275 Grade 11 students to review the questions within the measure for suitability to be included. The FriedBan Test Anxiety Scale also provides statistical validity and reliability data (see Table 1 above).

Across the range of the measures, wide-ranging and diverse underpinning constructs were observed, and these are reported in Table 2 below; these included expected aspects of worry and physiological factors, but also elements such as general school concerns, dread, and manifest dream anxiety, as well as a range of social concerns.

***Table 2: Underpinning constructs by proposed domain***

| Predisposition | Emotion | Behaviour | Cognitive | Physical | Social |
|---|---|---|---|---|---|
| Worry (Trait) | Worry (State) | Effort | Motivation | Physiological Symptoms | Social Comparison |
| General School Concerns | Fear | Off Task Behaviour | Negative Thoughts | Autonomic Reactions | Social Expectation |
| General Fears/Worries | General Fear | Behavioural Responses | Cognitive Impairment | Physiological Anxiety Responses | Social Status |
| | Anxiety | Task Irrelevant Behaviour | Cognitive Responses | Physiological Indicators | Social Desirability |
| | Manifest Anxiety | Inattention | Cognitive Interference | Bodily Symptoms | Social Concerns |
| | Dislike of School | Lack of Confidence | Test Irrelevant Thinking | Physical Tension/Excitement | Social Humiliation |
| | Emotionality | | Interference | Somatic Symptoms | |
| | Dread | | Poor Self-Evaluation Skills | Somatic Signs of Anxiety | |
| | Threat to Self-Image | | Cognitive Obstruction | Physiological Hyperarousal | |
| | Tension | | | | |
| | Restlessness | | | | |
| | Failure | | | | |
| | Emotional Excitement | | | | |
| | Test Anxiety | | | | |
| | Recitation Anxiety | | | | |
| | Manifest Dream Anxiety | | | | |

## Discussion.

### *Test anxiety measure construction and use.*

All the measures of test anxiety that were identified were found to have been developed using norm-referenced sample groups of varying sizes and age profiles; this being a common approach when developing a measure that aims to compare an individual's reported levels of test anxiety against that of a wider group or population (Hopkins, 1998). However, when taking this approach, it is important to ensure that any sample group fully represents the broader population with whom the measure will be used. This includes ensuring the representation of suitable age groups, balanced genders, and inclusion of ethnic and minority groups and individuals with special educational needs.

In wider research use, it is also important that any measure selected for use within a study fully reflects the broad community within which it is to be used to ensure the validity of any findings (Herrnstein & Murray, 1996). Given the finding that 14 of the 20 measures were developed with undergraduate students, this raises some concerns about the appropriateness of the norm sample group as, although undergraduate students provide an easy and cost-effective way to conduct research, their use in the development of measures assumes that this group is representative of the general population. Research evidence highlights concerns around the use of undergraduate sample groups as they are likely to be culturally and socially biased as they tend to make conclusions about human nature based on samples taken solely from predominantly Western undergraduate students who are more academically successful and educated to a higher level than the general population (Henrich, Heine & Norenzayan, 2010). Hanel & Vione (2016) highlight further concerns around generalising from student populations as they need to be configured regarding possible political bias when studying personal and attitudinal variables.

These findings also clearly provide evidence that some measures are being widely used

outside the countries where they were developed. As such, caution must be exercised when using any measure outside of the area where it was standardised, e.g., American measures being used in say, Germany. While the American population may share similarities with that of Germany in various aspects, they are not identical. For example, the German education system greatly differs from that in the Unites States of America in terms of its three-tier school structure, early specialisation, and longer duration. Given these differences, it is not reliably the case that the population on which the measure was constructed is equivalent to the group on which it was used, which inevitably impacts its applicability.

The age of the measures used within the body of research is also of concern. Herrnstein & Murray (1996) and James (2013) suggest that generally, a norm-referenced measure may be considered outdated if the data used to create the norm was collected more than five years ago. Clearly, norms tend to change over time due to social, cultural, and economic changes and the introduction of new academic research standards. Table 1 shows that over 90% of the studies we identified used a test anxiety measure based upon a norm sample group older than five years. This questions the validity of the obtained results as they may not necessarily represent the current population profile, potentially leading to questionable or inaccurate results that affect the validity of any conclusions that are drawn.

***Terms, concepts, and language used within test anxiety measures.***

Within the body of literature, there is some commonality around test anxiety being made up of four distinct aspects, these being: (i) worry, (ii) emotionality, (iii) cognitive interference, and (iv) physiological factors (cf. Putwain & Daly, 2014). However, having examined the language used in the published manuals or development papers for each measure (see Table 2), more than 40 different terms were identified that refered to a broad (and sometimes overlapping) range of aspects relating to test anxiety. The researcher grouped these terms into at least six distinct semantic domains: (i) Predisposition, (ii) Emotionality, (iii) Behaviours,

(iv) Cognitive, (v) Physical, and (vi) Social, and as such it is argued that this indicates that current research around test anxiety is predicated on a linguistic construct that is far more diverse than that suggested by Putwain & Daly (ibid).

It is suggested that the lack of consistency in terms, concepts, and language results from a lack of consensus around the test anxiety construct. As such, it is argued that the term test anxiety, can only be seen as a broad, portmanteau term that highlights the complexity of views, experiences and emotions that make up test anxiety. Moreover, given that self-reporting test anxiety measures have been developed using this variation in terms, which is open to individual interpretation, it calls into question the validity of some of the measures and the reliability of the evidence base.

As a result, it is argued that there is a need for greater consistency when discussing test anxiety for two principal reasons. Firstly, the wide range of terms, concepts, and language used across measures risks undermining the conceptual coherence and understanding of test anxiety, as researchers may not be discussing or measuring the same things. Secondly, consistency in the terms, language, and concepts will allow for more focused investigations, resulting in more comparable and generalisable findings.

***Lack of involvement of children and young people in developing test anxiety measures.***

Linked to the above, in the introduction to this paper, it is noted that there is a need for a person-centred approach and the inclusion of the voice of children and young people within educational practice. Given that these findings indicate that only the FriedBen Test Anxiety Scale (Friedman & Bendas-Jacob, 1997) referred to a development checking process that invited children and young people to review and provide feedback on the measure's construction. It is argued that this raises broader ontological concerns in that most test anxiety measures (and the research based on its application) utilise an a priori, adult-voiced construct

of test anxiety applied to children and young people.

Article 12 of the United Nations Declaration on the Rights of the Child states that children should be allowed to express their opinions and have their views considered in any matters affecting them. As such, it is argued that it is the correct ethical approach to involve children and young people in developing measures and additionally this will also help to ensure that a measure's content validity closely aligns with children and young people's contemporaneous experiences.

Howard (2020) suggests that a reliable measure of high-test anxiety is needed to support our understanding of test anxiety and its possible wide-reaching effects on children and young people's academic performance, self-confidence, and mental health. Working closely with children and young people facing important examinations, teachers, school leaders and educational psychologists are ideally placed to conduct this further research. However, given the concerns raised here about the age of some test anxiety measures, their sample group selection, as well as the lack of involvement of children and young people in the content/construct checking, none of the measures identified can be considered as being wholly satisfactory in identifying levels of test anxiety.

Selection of a suitable measure of test anxiety will depend upon several factors, such as the length of time it takes to administer the measure, the geographical location of the proposed research matching the nationalities of the measure's norm-group sample, the number of times a measure has been used within a particular country if comparisons are to be made with broader research, as well as the measure covering the particular constructs of test anxiety (e.g., worry, physiological symptoms, etc.) that need to be investigated in the research. Those wishing to conduct research within the UK secondary school population might consider the Multidimensional Test Anxiety Scale (MTAS) (Putwain et al, 2020). This measure covers the four widely recognised constructs conceived to make up test anxiety and has the most recent

norm sample group that includes secondary-aged students. However, children and young people do not appear to have been included within its development and as such it is not certain that the measure fully reflects their views of the test anxiety experience.

However, those working internationally and who wish to compare their findings with groups in other countries may wish to use one of the measures used across multiple countries. These include measures such as the Test Anxiety Inventory (Speilberger et al., 1970) or Test Anxiety Scale (Sarason, 1976). However, any research carried out with such older measures should highlight its provenance including the age of the measure and the ages of their sample-norm when reporting their results. Measures where validity or reliability data are currently unavailable, such as the Bakere Test Anxiety Scale (Bakare, 1969) and the Test Anxiety Questionnaire for Children (Donolato & Altoè, 2017), might best be avoided until such time that supporting data becomes available.

***Implications for future research around test anxiety.***

It is clear from the wide range of constructs used in measures of test anxiety that there is currently a lack of consensus on what exactly constitutes test anxiety. Furthermore, it is important to understand that children and young people will have unique perspectives on test anxiety and how it affects them, that may not fit within current perspectives and understandings of (adult) researchers. Given that most current instruments appear to have been developed using a priori constructs, there is a need for future research to design and validate measures where contemporary children and young people's views, experiences, and meanings around test anxiety are considered. This will help avoid misunderstanding children's and young people's experiences and better enable the link between support provision and needs.

These findings indicate that research on test anxiety has been conducted in 24 countries, with several papers being published in certain countries such as the USA, UK, and Germany.

However, notable absences in published research were observed in some comparable English-speaking countries such as Australia and New Zealand. It is possible that there is an inherent cultural bias within the outcomes of test anxiety research rather than in the development of the measures themselves. As such, further international research is needed to understand why some countries are represented well within the body of literature and why others are absent. The uneven geographic spread of countries originating research papers in this area provides an indication of *'pockets'* of specialist researchers as well as wider cultural differences around test anxiety that may be influenced by differing national incidence levels, health priorities, and the development of, or reach within, educational systems.

Examination of the measures also highlights aspects of social concerns, including comparisons to an individual's wider peer group, social status, and social humiliation. Bozzola et al. (2022) highlight a year-on-year increase that accelerated sharply during the COVID-19 pandemic in children's access to and use of online social media platforms and a possible link between social media use and unfavourable peer comparison and social humiliation. Given that all the test anxiety measures identified within Table 1 were developed before the pandemic, it is not known whether these social aspects of test anxiety are now a more prevalent concern for children and young people. As such, further research in this area is needed.

**Limitations.**

We have identified three possible limitations to the research reported here. First, we have deliberately excluded papers outside the English language as one of the search criteria. As a result of this decision, this may have excluded papers that referenced the use of a test anxiety measure, and this will not have been included within this study. However, we did identify a range of international papers that identified a range of measures that were developed across the 24 countries highlighted in Figure 2, including the USA, Germany, Israel, UK, and India.

The second limitation is that, arguably, the second research question is predicated on the underlying ontological position that the views of children and young people *should* be listened to and incorporated into the development of any measure that seeks to identify their test-related emotions. This assumption may be seen as a form of cultural bias since not everyone views such inclusion as necessary. While there is a position that the development of psychological measures is the business of psychological professionals, the researcher would point in response to the increasing and worldwide acceptance of the 1994 Salamanca Statement (UNESCO, 1994) that sets out that the educational systems and programmes should be inclusive in design and delivery. Therefore, this supports the assertion that this ethical approach is suitable, and involving engagement and co-construction will help ensure that a measure's content validity closely matches the current experiences of children and young people.

The third limitation is that we have taken the absence of any reference to children and young people being engaged in the development process of a test anxiety measure as being evidence that there has not been any involvement – these findings indicate that just one measure (FriedBen Test Anxiety Scale, Friedman & Bendas-Jacob (1997) directly referenced the involvement children and young people during its development process. It is possible that other measures were developed with children and young people but that this has not been mentioned by researchers, possibly due to them having distinct research priorities or due to word limitations imposed by the publication process.

**Disclosure statement**

The author reported no potential conflicts of interest.

**References**


Ajilchi, B., & Nejati, V. (2017). Executive functions in students with depression anxiety, and stress symptoms. *Basic Clinical Neuroscience*, 8(3), 223-232.

Bakare C.G.M., (1969). Phenomenal Self- Concept, Anxiety and Academic Performance. Unpublished PhD. Thesis, Columbia University.

Benson, J., El-Zahhar, N. (1994). Refinement and validation of the revised test anxiety Questionnaire. *Structural Equation Modelling*, 1 (3), 203-221

Bozzola E., Spina G., Agostiniani R., Barni S., Russo R., Scarpato E., Di Mauro A., Di Stefano A.V., Caruso C., Corsello G., Staiano A. (2022) The use of social media in children and adolescents: Scoping review on the potential risks. *International Journal of Environmental Research in Public Health*. 19(16), 9960.

Butler, R. J., & Gasson, S. L. (2005). Self-esteem/self-concept scales for children and adolescents: A review. *Child and Adolescent Mental Health*, *10*, 190-201.

Byrne, B. M. (2002). Validating the measurement and structure of self-concept: Snapshots of past, present and future research. *American Psychologist*, *57*, 896-909.

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

Cizek, G. J., & Burg, S. S. (2006). *Addressing test anxiety in a high-stakes environment: Strategies for classroom and schools.* Corwin Press.

Cohen, L., Manion, L. & Morrision, K. (2011). *Research methods in education.* London: Routledge.

Damer, D & Melendres, L. (2011) Tackling test anxiety: A group for college students. *Journal for Specialists in Group Work*, 36:3, 163-177

Donolato, E., Marci, T., Altoè, G., & Mammarella, I. C. (2020). Measuring test anxiety in primary and middle school children: Psychometric evaluation of the Test Anxiety Questionnaire for Children (TAQ-C). *European Journal of Psychological Assessment, 36*(5), 839–851.

Driscoll, R. (2006) Westside test anxiety scale validation. Education Resources Information Centre.

Eysenck, H.J., & Rachman, S., (1965) The Causes and Cures of Neurosis: An Introduction to Modern Behaviour Therapy based on Learning Theory and the Principles of Conditioning, London: Routledge

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

Harter, S. (1999). *The construction of the self: A developmental perspective*. New York, NY: Guilford Press.

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

Hopkins, K.D. (1998). *Educational and psychological measurement and evaluation.* New York: Allyn & Bacon

Howard, E. (2020). *A review of the literature concerning anxiety for educational assessme*nts. Ofqual: Coventry.

James, T. (2013). Educational assessment: A study guide for the Certificate of Competence in Educational Testing.

Jones, S. & Forshaw, M (2012). *Research Methods in Psychology*, London: Prentice Hall

Keogh, E., Bond, F. W., & Flaxman, P. E. (2006). Improving academic performance and mental health through a stress management intervention: Outcomes and mediators of change. *Behaviour Research and Therapy*, 44(3), 339-357.

Lapouse, R., & Monk, M. A. (1959). Fears and worries in a representative sample of children. *American Journal of Orthopsychiatry*, 29(4), 803–818.

Lowe, P. A. (2014). The Test Anxiety Measure for Adolescents (TAMA): Examination of the Reliability and Validity of the Scores of a New Multidimensional Measure of Test Anxiety for Middle and High School Students. *Journal of Psychoeducational Assessment*, *32*(5), 404–416.

Lowe, P. A., Grumbein, M. J., & Raad, J. M. (2011). Examination of the Psychometric Properties of the Test Anxiety Scale for Elementary Students (TAS-E) Scores. *Journal of Psychoeducational Assessment*, *29*(6), 503–514.

Lowe, P. A., Lee, S. W., Witteborg, K. M., Prichard, K. W., Luhr, M. E., Cullinan, C. M., 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.

Mccaldin, T. (2020) The experiences of students preparing for GCSE examinations: An interpretative phenomenological analysis. Unpublished Phd Thesis, University of Manchester.

Mandler, G., & Sarason, I. G. (1952). A study of anxiety and learning, *Journal of Abnormal and Social Psychology*, 87, 166-173.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Open Medicine, 3(3), 123-130.

Pekrun, R., Muis, K.R., Frenzel, A. C., & Goetz, T. (2018). *Emotional at School*. New York: Routledge.

Petermann, F., Winkel, S. (2017). FLM 7-13 Fragebogen zur Leistungsmotivation für Schüler der 7. bis 13. Klasse. Bern: Hogrefe.

Peleg, O. (2009). Test anxiety, academic achievement and self- esteem among Arab adolescents with and without learning disabilities. *Learning Disabilities Quarterly*, 32, 11-20.

Putwain, D., & Daly, A. L. (2014). Test anxiety prevalence and gender differences in a sample of English secondary school students. *Educational Studies, 40*(5), 554–570.

Putwain, D.W., Daly, A.L., Chamberlain, S., & Sadreddini S. (2015). Academically buoyant students are less anxious about and perform better in high-stakes examinations. *British Journal of Educational Psychology*. 2015 Sep;85(3):247-63.

Putwain, D. W., von der Embse, N. P., Rainbird, E. C., & West, G. (2020). The development and validation of a new Multidimensional Test Anxiety Scale (MTAS). *European Journal of Psychological Assessment*. Advance online publication.

Robson, C. (2011). *Real world research* (3rd ed.). Chichester, UK: Wiley.

Rodway, C., Tham, S.G., Ibrahim, S., Turnbull, P., Windfuhr, K., Shaw, J., Kapur, N., Appleby, L., (2016) Suicide in children and young people in England: A consecutive case series. *Lancet Psychiatry*. 3(8), 751-759.

Segool, N.K., Von der Embse, N., Mata, A.D, & Gallant, J. (2014) Cognitive Behavioral Model of Test Anxiety in a High-Stakes Context: An Exploratory Study. *School Mental Health*, 6, 50-61

Seipp, B. (1991). Anxiety and academic performance: A meta-analysis of findings. *Anxiety Research*, 4(1), 27–41

Spielberger, C. D. (1980). *Preliminary Professional Manual for the Test Anxiety Inventory.* Palo Alto, CA: Consulting Psychologists Press.

Spielberger, C. D., & Vagg, P. R. (Eds.). (1995). *Test anxiety: Theory, assessment, and treatment.* New York: Taylor & Francis.

Spielberger, C. D., Gorsuch, R. L., Lushene, R., Vagg, P. R., & Jacobs, G. A. (1983). *Manual for the State-Trait Anxiety Inventory*. Palo Alto, CA: Consulting Psychologists Press.

Sung, Y.-T., & Chao, T.-Y. (2015). Construction of the Examination Stress Scale for Adolescent Students. *Measurement and Evaluation in Counselling and Development*, *48*(1), 44–58.

Tatlow-Golden, M., Guerin, S. (2017) Who I am: The meaning of early adolescents most valued activities and relationships, and implications for self-concept research. *Journal of Early Adolescence*. 37(2) 236–266

Thomas, C. R. and Gadbois, S. A. (2007), Academic self-handicapping: The role of self-concept clarity and students' learning strategies. *British Journal of Educational Psychology*, 77: 101–119.

Thomas, C. L., Cassady, J. C., & Finch, W. H. (2016). Identifying severity standards on the cognitive test anxiety scale: Cut score determination using latent class and cluster analysis. *Journal of Psychoeducational Assessment*. 37(2) 226–232

Ursache, A., Raver, C.C., (2014) Trait and state anxiety: relations to executive functioning in an at-risk sample. *Cognition & Emotions*. 28(5):845-55.

Von der Embse, N., Barterian, J. and Segool, N. (2013) Test Anxiety Interventions for Children and Adolescents: A Systematic Review of Treatment Studies from 2000–2010. *Psychological Schools Journal*. 50: 57–71.

Wieczerkowski, W., Nickel, H., Janowski, A., Fittkau, B. & Rauer, W. (1975). *Angstfragebogen für Schüler* (*AFS).* Westermann, Braunschweig.

Wren, D. G., & Benson, J. (2004). Measuring test anxiety in children: Scale development and internal construct validation. *Anxiety, Stress & Coping: An International Journal, 17*(3), 227–240.

Zeidner, M. (1998). Perspectives on individual differences. Test anxiety: The state of the art. New York: Plenum Press.

Zeidner, M., & Matthews, G. (2005). Evaluation Anxiety: Current Theory and Research. In A. J. Elliot & C. S. Dweck (Eds.), Handbook of competence and motivation (141–163).

# Paper 2: Testing Times: A Longitudinal Study of Test-Related Emotions

Paper prepared in accordance with the author guideline and formatting template for the journal, Applied Measurement in Education (See Appendices 1 & 2).

Word count: 9103 (excluding references and tables)

## Abstract

Testing and assessments are integral to educational systems designed to evaluate students' knowledge and abilities across subject areas. However, research evidence indicates that testing could result in emotional challenges affecting student performance and long-term mental health. Test anxiety can lead to attentional and cognitive processing difficulties, negative thoughts, low self-worth and can harm physical health. However, the reported prevalence of high levels of test anxiety varies, and it is often unrecognised by teachers and school leaders.

The findings indicate that test emotions can vary over time and that they appear to be influenced by contextual factors in the school setting. Test Enjoyment, Test Hope, and Test Anger remain relatively stable throughout the school year. In contrast, Test Relief, Test Anxiety, Test Shame, and Test Hopelessness fluctuate in response to significant testing events such as mock and final examinations. We also found differences in reported test emotions between schools, with students from School A reporting lower levels of Test Pride, Test Relief, and Test Anxiety and higher levels of Test Shame and Test Hopelessness than their School B peers. The findings suggest that professionals working with young people could support their emotional preparedness for test experiences by proactively addressing test emotions and considering the dynamic nature of those emotions throughout the school year.

Future research opportunities include understanding the impact of school culture on test emotions, investigating individual and group differences in test emotions, and identifying optimal timings for interventions.

Keywords**:** test emotions, academic year, schools, examination stress

## Introduction

Testing and assessments are integral components of educational systems that are designed to evaluate students' knowledge and abilities across subject areas (Stobart, 2008). However, Pekrun et al. (2018) points to the emotional challenges that testing, and examinations create, highlighting their potential to undermine student performance as well as the possible long-term consequences for children and young peoples' mental health and academic achievement. Pekrun (ibid.) believes that emotions aroused through the testing process are complex, multifaceted experiences that are shaped by a range of individual, social, and contextual factors and argues that they play a crucial role in students' academic performance. Consequently, understanding the nature of these emotions is central to improving educational outcomes.

### *The impact of test emotions on young people.*

Although some find the heightened emotions around testing to be facilitative, wider research indicates that high levels of anxiety in children are associated with attentional and cognitive processing difficulties that can have a detrimental impact on academic performance (Hembree 1988; Seipp 1991; Cassady & Johnson, 2002; Keogh et al. 2004; Ursache & Raver, 2014; Ajilchi & Nejati, 2017). The more severe the anxiety, the greater the impact on academic performance. This is particularly relevant for students who experience high levels of anxiety in testing situations, as this can have an adverse effect on their ability to demonstrate their knowledge and skills during assessments.

Furthermore, research has highlighted that high levels of anxiety focused on testing and assessment can negatively influence self-esteem (Peleg, 2009; Thomas & Gadbois, 2007). Children and young people who experience test anxiety often have negative thoughts and beliefs about their abilities, which can lead to feelings of inadequacy and low self-worth (Thomas & Gadbois, 2007). These negative emotions can be particularly harmful in academic

settings, where children are constantly evaluated and compared to their peers (Stobart, 2008). Test anxiety can exacerbate these feelings and make it harder for children to maintain a positive self-image and a growth mindset.

In addition to the academic impact, long-term exposure to higher levels of anxiety can have detrimental effects on physical and mental health. Studies have shown that anxiety can disrupt sleep patterns, increase the risk of mental health difficulties, and suppress the immune system, making individuals more susceptible to various illnesses (Damer & Melendres, 2011; Mental Health Foundation, 2014). The implications of these findings are significant, as they suggest that untreated anxiety in childhood can have long-lasting effects on individuals' health and well-being.

Perhaps the most concerning finding is that high levels of examination pressures have been implicated as a significant antecedent factor in incidents of self-harm and suicide among adolescents (Rodway et al., 2016). The intense pressure to perform well in exams and assessments can be overwhelming for some students, leading to feelings of hopelessness and despair. This highlights the urgent need to address test anxiety in schools and provide appropriate support and interventions for children and young people who are struggling with these emotions.

### *Prevalence of test anxiety*

While we have an understanding around impact of test anxiety, reported levels of prevalence vary greatly across research studies. Factors that contribute to the variation in reported prevalence include a lack of routine screening, the reliability of testing measures, and differing thresholds for reliable differentiation. A significant body of research suggests that test anxiety may affect up to 40% of the school-age population, while recent UK-based studies indicate that up to 16% of students in Year 11 may experience high levels of test anxiety (von der Embse, Barterian & Segool, 2013).

Putwain & Daly (2014) highlight that despite the prevalence of test anxiety, it remains largely unrecognised by headteachers as a current strategic priority and suggesting that they often do not identify test anxiety and exam stress as a significantly detrimental factor to their school's examination results. Additionally, they point out that the high TA prevalence rate of 16.4% of GCSE students reported in studies such as studies was not recognised by headteachers as reflective of their schools' typical Year 11 cohorts. In their experience, headteachers suggested a much lower level of need within school populations, with just two or three students per year struggling with significant levels of test anxiety.

***Patterns of test emotions***

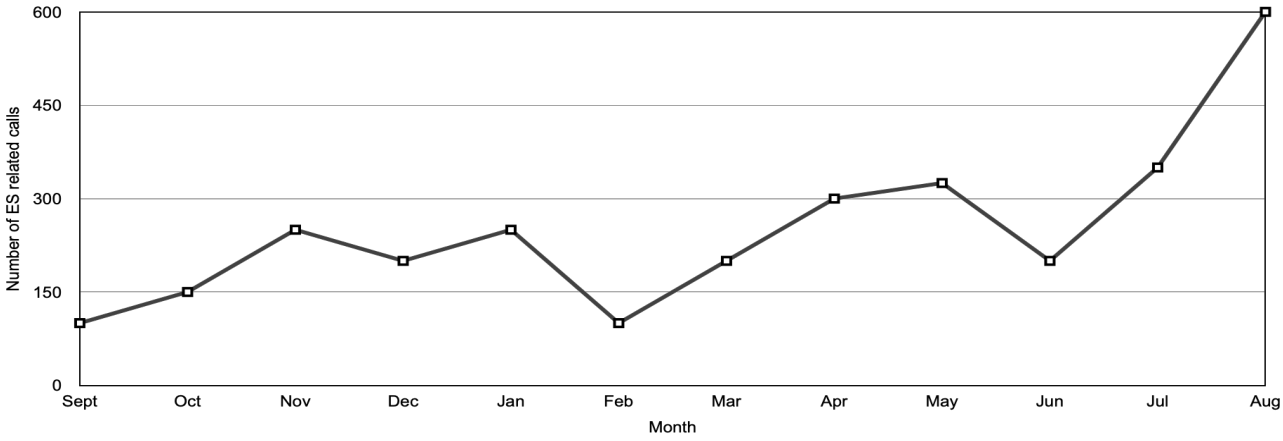
Linked to the above, Pekrun et al. (2018) highlights the need for a clearer understanding of young people's test emotions including the prevalence of test anxiety to inform the development of interventions and changes in schools' practice around supporting children and young people who experience high levels of test anxiety. Further, Pekrun et al. (2018) emphasises the importance of understanding the patterns of children's test emotion over time, highlighting the need to differentiate between trait test anxiety, which is a general level of anxiety that individuals possess, and state test anxiety, which can fluctuate depending on the circumstances.

However, it is important to note that test emotions may not necessarily be consistent and static, rather, it is possible that they can vary over time in response to factors such as increased academic pressures that students face as they progress through the academic year (Pekrun, Elliot, and Maier, 2009) and individual contextual factors, such as previous academic achievement, cognitive ability, and personality traits (Schneider & Preckel, 2017).

This possible variation is illustrated through data provided by the NSPCC's ChildLine service showing the number of calls received from children and young people concerning examination stress. Figure 1 (below) illustrates a pattern of gradually increasing call numbers

across the academic year, with up to one-third of their exam-related counselling sessions occurring in the first five months of the academic year, well in advance of the May/June exam season.

***Figure 1: Calls to the ChildLine service linked to exam stress (Bentley et al., 2019)***




***Variation in test emotions across settings.***

However, while there is a body of research which indicates possible school contextual factors that may influence children's and young people emotions around testing and examinations, there is a gap in knowledge around which emotions and how these may differ across different timepoints.

For example, Wang, Eccles, & Kenny (2013) point to factors such as school culture (e.g., the level of competition, academic expectations, the value placed on test scores) as being influencing factors that have the potential to create a more stressful environment for students. They go on to suggest that this pressure to perform well can trigger anxieties, especially in students who are already predisposed to more generalised anxiety.

According to Mccaldin (2020), it is important to note that fear appeals, which involve teachers communicating the potential outcomes of success and failure to students, can also impact the emotional state of young individuals. Linked to this, Li and Lai (2011) found that authoritarian teaching practices were associated with higher levels of test anxiety. Similarly,

Salend and Garrick Duhaney (2005) found that students who perceived their teachers as using a more authoritarian teaching style reported higher levels of test anxiety. Research has also shown that focusing solely on exam preparation can contribute to test anxiety. For example, a study by Hembree (1988) found that students who received more test preparation instruction reported higher levels of test anxiety, and Putwain and Daly (2013) found that students who perceived their teachers as placing high emphasis on testing reported higher levels of test anxiety.

## Study Aims and Research Questions

Given the current position outlined above, it is argued that single, one-time measures of test anxiety are insufficient to fully investigate and identify possible trends and patterns across the time. Research that tracks the changes in children's test emotions throughout the school year is more comprehensive and enables better understanding of how children and young people experience test emotions and effective strategies to support them.

The aims of this study are to investigate the pattern of test emotions and how they vary over time, as well as to identify contextual factors within the school setting that may influence test emotions. The findings are intended to inform the development of effective strategies to address test anxiety and other test-related emotions in students, and to guide post-examination provision for children and young people. Specifically, the proposed research questions are:

(1) How do test emotions change across the academic year?

(2) Do these test emotions differ across school settings?

## Method

### *Research Design*

This research was a survey design, employing a quantitative approach to data collection in

line with a relativist ontological position and a positivist epistemological stance. While the overall endeavour of this paper employs a quantitative, closed measure, which could be perceived as leaning towards a realist position, the interpretation of the data and the overall intent of the research demonstrate that the researcher is not attempting to promote a simplified, universally applicable truth that applies to all schools.

This approach aligns with Muijs (2011) and Cohen, Manion, and Morrison (2011), who highlight this approach for the use of quantitative, objective data in research to identify patterns and establish causal relationships.

### *Participants*

As recommended by Patton (2015), a criterion sampling approach was selected to ensure the homogeneity of the sample population based on the pre-determined criterion of Year 11 students in secondary state school settings. Special schools were excluded from this study as examination arrangements in these settings can be very different from mainstream schools. Two participating schools were recruited from within the same local authority in England although these were in contrasting locations as described in Table 1 below.

## *Table 1: Details of school settings* (ONS, 2019)

| School Context | School A | School B |
|---|---|---|
| | The school is located within a large residential estate that is predominantly made up of council and social housing, resulting in lower-than-average levels of home ownership. Historically, the area has had relatively low socio-economic status, with high levels of poverty and deprivation. This has been reflected in government statistics, which rank the area in the top 20% of the most deprived areas in England. The area has higher rates of unemployment, low income, and low educational attainment compared to national averages. | The school operates within a relatively affluent socio-economic environment. The town has a mix of residential areas, including some high-value properties, which contribute to the overall affluence of the area. The area has a mix of owner-occupied homes and rental properties, with a higher-than-average proportion of residents owning their homes. The area has a relatively low unemployment rate, and there are job opportunities available in a variety of sectors, including manufacturing, retail, and professional services. |
| Index of Multiple Deprivation Decile | 2 | 4 |
| The Education & Skills Decile | 1 | 4 |
| Income Decile | 2 | 4 |
| Employment Decile | 2 | 5 |
| Average house price 2021/2022 | £202,532 | £424,182 |
| School Ofsted Rating | Good | Outstanding |
| Number on roll | 634 | 721 |

### *Instrument*

For data collection, the researchers used the Test-Related Emotion Scales developed by Pekrun et al. (2004), which provided a reliable measure of a range of emotions. The scale provides a measure of a range of test emotions including enjoyment, hope, pride, relief, anger, anxiety, shame, and hopelessness. Overall test-retest reliability for the measure is given by Pekrun et al. (ibid) as being, r = .65 and r = .86. The authors also report a validity coefficient of r = 0.80 in correlation with the Test Anxiety Inventory (TAI).

The data was collected through an online questionnaire administered on the Qualtrics platform (See Appendix 5). Participants were asked about their typical feelings before, during, and after taking a test or exam over the previous eight-week period, using a Likert scale of 1-5, where “1” indicates "*Not like me*", “2” indicates “*A little like me*”, “3” indicates “*Sometimes like me*”, ‘4” indicates “*Often like me*”, and “5” indicates "*Very much like me*". To make the questionnaire convenient for participants to complete on a smartphone, slider bars were used for responses instead of numeric inputs, as they were easier and quicker to use on mobile devices.

### *Data Analysis Method*

The data was collated and transferred into the IBM SPSS Statistics (Version 27) for analysis. The data was screened prior analysis to ensure its quality, identify outliers, check for normality and assumptions, and determine appropriate analysis methods. Given that participants may have not provided data at each time point due to factors such as school absence, the dataset was examined using Little's MCAR test to determine if missing data were missing completely at random (MCAR) or not. The analysis indicated that there was no relationship between the missingness of the data and any values, observed or missing: ($\chi^2(852) = 633.950$, $p = 1.00$). As such, the missing data points appear to be a random subset

of the data. Identifying whether missing data is missing completely at random (MCAR) is important because the missingness mechanism can affect the validity and accuracy of statistical inferences drawn from a dataset.

Prior to further analysis, it was important to also determine the sampling adequacy of the data collected. G*Power was again used with a medium Cohen F effect size of f=0.25, power (1-α) = 0.95, non-sphericity correction of ε = 1, 6 measurements and five groups, that is two from school location and three from gender. The results show that the non-central parameter was λ = 32.81, $F_{crit}(20, 325) = 1.60$ and the corresponding minimum sample size required for each group was 70. All the groups met this criterion except for the gender category *other* which had only 3 responses. There were 79 males, 86 females, 88 School A students and 80 School B students; therefore, a 2x2 repeated measures factorial design was used.

Bartolucci, Bacci, and Gnadi (2016) and Hair et al. (2019) suggest that a repeated measures factorial analysis design is the best way to analyse the data in a study of this design. This is because the study involves multiple measures of the same variables, rated by the same students over six different time periods. To do this, the researchers examined the data for each of the eight test emotions and the four wider, composite emotions to establish overall means and ranges at each time point. A two-way Factorial Repeated Measures ANOVA was also carried out to determine if there were any statistically significant differences in the emotions over time.

### *Ethics*

The research was carried out in line with the General Data Protection Regulations and Data Protection Act (GDPR) as stated in the University of Manchester guidelines and information governance policies (University of Manchester, 2020). It also adhered to the Ethical Practice Policy and Guidance established by the Manchester Institute of Education and The British Psychological Society's Code of Human Research Ethics (2014). It underwent review and

approval by the University of Manchester Ethics Committee (UREC), with the assigned reference number: 2021-12798-20226. Additionally, to ensure continuous oversight, a professor specialising in educational psychology and an experienced professional in educational research were designated as supervisors, available to provide ongoing support and address any concerns that may arise during the research process.

Before participation, informed consent was obtained, and the purpose of the research was communicated through school assemblies and/or PSHE sessions. All participants were provided with a Participant Information Sheet at the start of the Qualtrics survey (Appendix 1). This explained exactly what the research project was focused on before seeking consent to participate in the research. Participants were informed that they could withdraw from the research at any time before and during the survey. Participants were also informed that they could not have their data removed once submitted as it was not possible to identify individual responses.

Participants were not required to share the identities of any family members, peers, or professionals with whom they had worked. In addition, school names were not included to minimise the possibility of identification.

As the survey asked questions about the participants' feelings, which they may have found difficult, a destress protocol was drawn up and appropriate levels of support were pointed to on submission of the survey. This ranged from initially discussing their concerns with parents or a teacher with whom they felt comfortable, through to contact details for Childline.

### *Quality & Rigour*

The researcher implemented several quality assurance measures to ensure the credibility of the claims and findings derived from the research:

(1) **Confirmability audit:** The original raw data, including the responses collected from the online questionnaire, were preserved, and are available for scrutiny or re-analysis by others, if required. This establishes a "chain of evidence" as recommended by Yin (2013), allowing for transparency and verification of the data.

(2) **Dependability audit:** A diary was maintained throughout the research process, documenting each step, including data collection, data analysis, and interpretation of results. This diary provided a record of the research process, facilitating an audit of the research steps and ensuring consistency between the research design and the research questions, as suggested by Miles and Huberman (1994).

(3) **Reliability of data:** The data collected was collated and shared with the research supervisors, who reviewed the data set for accuracy and consistency. To assess the internal consistency of the data for parametric analysis, Cronbach's alpha was utilised. The resulting analysis revealed that the Cronbach's alpha values for the data ranged from 0.61 to 0.82, with an average of 0.72.

These quality assurance measures aimed to enhance the credibility of the research findings by ensuring transparency, consistency, and reliability of the data and research process. By making the raw data and research process available for scrutiny by others, the researcher aimed to establish the credibility of the research and provide a foundation for valid and reliable conclusions.

**Findings**

The findings of the study revealed statistically significant differences in the mean ratings of all eight test emotions across the six time periods. There were no interaction effects with gender, indicating that gender did not significantly influence the test emotions reported by the participants.

However, there were significant interactions between school type and some of the test emotions. Specifically, students from School A reported lower levels of test enjoyment, hope, pride, relief, and anxiety, and higher levels of test anger, shame, and hopelessness compared to their peers in School B.

Since there were no established benchmarks for the Test-Related Emotions Scales (Pekrun et al., 2004), the researchers devised a rubric outlined in Table 2. This rubric was designed to ensure consistent descriptions of changes in test-related emotions over time. It was constructed based on the five-point Likert scale used within the questionnaire, with specific cut-off points calculated as one third of the total scale range. To clarify, the lower third of the scale range was classified as "low," the middle third as "average," and the highest third as "high". By employing this rubric, researchers were able to establish a uniform framework for interpreting and comprehending the reported changes in test-related emotions among the participants.

***Table 2: Rubric of Test Emotions Ranges***

| | Low | Average | High |
|---|---|---|---|
| Enjoyment (10 items) | <23 | 24 - 38 | >39 |
| Hope (8 items) | <19 | 20 – 30 | >31 |
| Pride (10 items) | <23 | 24 - 38 | >39 |
| Relief (6 items) | <14 | 15 - 23 | >24 |
| Anger (10 items) | <23 | 24 - 38 | >39 |
| Anxiety (12 items) | <28 | 29 - 45 | >46 |
| Shame (10 items) | <23 | 24 - 38 | >39 |
| Hopelessness (11 items) | <26 | 27 - 41 | >42 |
| Affective Emotions (21 items) | <49 | 50 - 78 | >79 |
| Cognitive Emotions (24 items) | <56 | 57 - 89 | >90 |
| Motivational Emotions (17 items) | <40 | 41 - 63 | >64 |
| Physiological Emotions (15 items) | <35 | 36 - 56 | >57 |

Table 3, below, presents a summary of the data collected from the survey for each of the test emotions: enjoyment, hope, pride, relief, anger, anxiety, shame, and hopelessness. For each

emotion, the table provides the mean and range of scores obtained at each time point. The range analysis indicates the spread of scores, with low ranges indicating that most scores fell close to the mean, while high ranges suggest greater variability in the data. The overall mean of the scores for each emotion are also provided, which gives a summary of the average score across all six time points. The mean score can be used to compare how the participants felt about each emotion on average throughout the survey. Figure 2 (below) presents a visual representation of the mean reported test emotions ratings over the six data collection points.

***Table 3: Mean Reported Test Emotions Ratings (All Participants)***

| | Enjoyment Test | | Test Hope | | Test Pride | | Mean Test Relief | | TestAnger | | Test Anxiety | | Test Shame | | Test Hopelessness | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Mean | Range | Mean | Range | Mean | Range | Mean | Range | Mean | Range | Mean | Range | Mean | Range | Mean | Range |
| T1 (Sept) | 23.81 | Low | 21.44 | Average | 27.55 | Low | 21.70 | Average | 23.07 | Low | 40.67 | Average | 21.47 | Low | 24.93 | Low |
| T2 (Nov) | 22.43 | Low | 21.95 | Average | 28.38 | Low | 21.23 | Average | 23.36 | Low | 41.93 | Average | 20.62 | Low | 24.46 | Low |
| T3 (Jan) | 20.71 | Low | 20.78 | Average | 28.46 | Low | 13.56 | Low | 24.59 | Low | 40.60 | Average | 23.84 | Low | 27.05 | Average |
| T4 (Mar) | 18.72 | Low | 22.34 | Average | 29.01 | Low | 20.94 | Average | 23.87 | Low | 47.97 | High | 21.56 | Low | 28.10 | Average |
| T5 (May) | 17.13 | Low | 21.78 | Average | 27.75 | Low | 17.78 | Low | 24.76 | Low | 47.98 | High | 28.86 | Average | 30.25 | Average |
| T6 (Jun) | 19.39 | Low | 21.55 | Average | 28.06 | Low | 20.89 | Average | 21.58 | Low | 41.91 | Average | 24.29 | Low | 25.69 | Low |
| Overall Mean | 20.32 | Low | 21.64 | Average | 28.20 | Low | 19.35 | Average | 23.54 | Low | 42.51 | High | 23.29 | Low | 26.69 | Low |

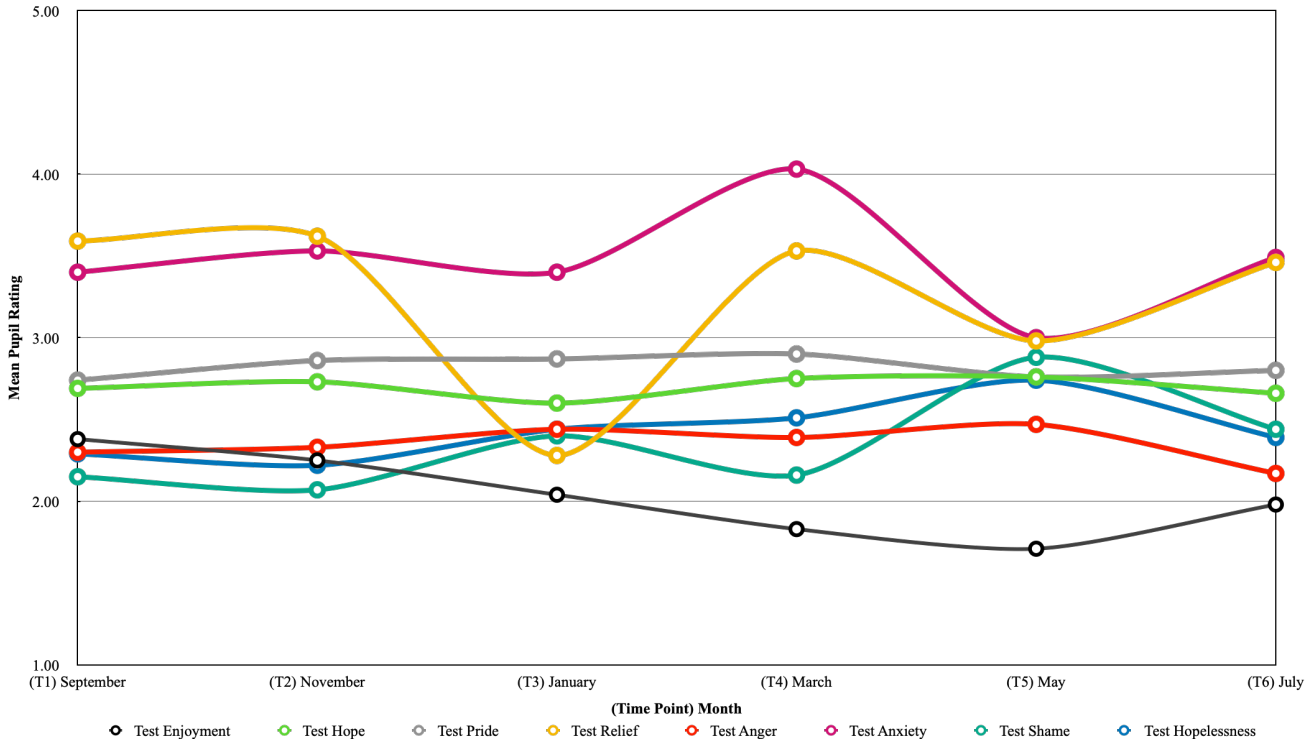
***Table 4: Response Rate of Participants***

| | Sept | | Nov | | Jan | | Mar | | May | | Jun | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Number | % | Number | % | Number | % | Number | % | Number | % | Number | % |
| School A | 65 | 90 | 68 | 94 | 54 | 75 | 60 | 83 | 58 | 81 | 55 | 76 |
| School B | 78 | 95 | 72 | 88 | 62 | 76 | 68 | 83 | 64 | 78 | 72 | 88 |
| Total | 143 | 93 | 140 | 91 | 116 | 75 | 128 | 83 | 122 | 79 | 127 | 82 |

## Response Rate

Table 4, above, provides details of participant response rates across the study. Overall, the mean response rate was 129 pupils at each time point, equating to 83.8% of the Year 11 population across both Schools A & B.

***Figure 2: Mean Reported Test Emotions Ratings (All Participants)***



### *Test Enjoyment*

The Test Enjoyment Scale provides a measure of the level of pleasure or satisfaction an individual experiences before, during, and after taking a test.

***Table 5: Mean Ratings – Test Enjoyment***

| | | | | Range Analysis | | |
|---|---|---|---|---|---|---|
| | Overall Mean | School A Mean | School B Mean | Low Range *Number (%)* | Average Range *Number (%)* | High Range *Number (%)* |
| T1 (Sept) | 23.81 | 22.69 | 24.94 | 77 (45%) | 91 (55%) | 0 (0%) |
| T2 (Nov) | 22.43 | 21.99 | 22.87 | 102 (70%) | 41 (30%) | 0 (0%) |
| T3 (Jan) | 20.71 | 21.22 | 20.21 | 125 (92%) | 11 (8%) | 0 (0%) |
| T4 (Mar) | 18.72 | 19.12 | 18.32 | 124 (92%) | 11 (8%) | 0 (0%) |
| T5 (May) | 17.13 | 17.33 | 16.94 | 148 (97%) | 8 (3%) | 0 (0%) |
| T6 (Jun) | 19.39 | 19.13 | 19.65 | 129 (81%) | 31 (19%) | 0 (0%) |
| Overall Mean | 20.32 | 20.25 | 20.49 | 118 (80%) | 32 (20%) | 0 (0%) |

The means of test enjoyment at each time point consistently fell within the low range across the school year. The highest mean score was observed in September ($M_{T1}$=23.32), indicating relatively higher levels of test enjoyment at the beginning of the school year. However, the reported levels of test enjoyment gradually declined over the following four time points, reaching the lowest point in May ($M_{T5}$=17.13), which coincided with the beginning of the examination season. After this low point, there was a slight recovery in reported levels of test enjoyment in June ($M_{T6}$=19.39), which coincided with the end of the examinations. None of the participants reported no level of test enjoyment.

The results of the range analysis presented in Table 5 indicate that throughout the school year, no individuals reported levels of test enjoyment that fell within the high range. At T1 (September), approximately 55% (n=91) of participants reported scores that fell within the average range. However, by T5 (May), this percentage drastically decreased to just 3% (n=6), with the remaining 97% (n=148) reporting low levels of test enjoyment. This shift in responses over time indicates that

individuals who previously reported average levels of test enjoyment shifted to indicate low levels of test enjoyment.

Furthermore, the published mean for test enjoyment in the norm group sample reported by Pekrun et al. (2004) was M=28.33 (SD=6.00), while the overall levels of test enjoyment found in this study were more than 1 standard deviation below the published mean. This suggests that participants in this study reported lower levels of test enjoyment compared to the norm group in Pekrun et al.'s (2004) study, indicating that young people in this study may have enjoyed testing less than those in the previous study.

The results of the statistical analysis indicate that there were statistically significant differences in the mean ratings of test enjoyment across the six time periods, with a significant F-value of (4.322, 375.976) = 41.690, $p<0.05$, and a moderate effect size of $\eta 2=0.324$.

Furthermore, there was a statistically significant interaction effect between test enjoyment and type of school, with a significant F-value of (4.322, 375.976) = 2.591, $p<0.05$, and a small effect size of $\eta 2=0.029$. Tukey's Least Significant Difference analysis revealed that this interaction effect was observed only at T1 (September), where students at the School A reported lower levels of test enjoyment compared to their counterparts in the School B.

This indicates that there were significant differences in the levels of test enjoyment across the six time periods, and that the school setting also played a role in influencing test enjoyment levels, particularly at the beginning of the school year (T1). Further analysis using post-hoc tests such as Tukey's LSD can help to understand the specific nature of these differences and provide more insights into the findings, however given word limitations these will be available in a forthcoming published research report.

### *Test Hope*

Test hope refers to an individual's level of optimism and confidence in their ability to perform well on a test. It reflects their belief in their own academic success despite challenges they may encounter.

### *Table 6: Mean Ratings – Test Hope*

| | | | | Range Analysis | | |
|---|---|---|---|---|---|---|
| | Overall Mean | School A Mean | School B Mean | Low Range *Number (%)* | Average Range *Number (%)* | High Range *Number (%)* |
| T1 (Sept) | 21.44 | 20.37 | 22.53 | 44 (26%) | 124 (74%) | 0 (0%) |
| T2 (Nov) | 21.95 | 21.75 | 22.05 | 23 (16%) | 120 (84%) | 0 (0%) |
| T3 (Jan) | 20.78 | 20.58 | 20.99 | 44 (32%) | 92 (68%) | 0 (0%) |
| T4 (Mar) | 22.34 | 21.75 | 22.49 | 35 (26%) | 94 (70%) | 6 (4%) |
| T5 (May) | 21.78 | 21.89 | 22.09 | 21 (14%) | 133 (86%) | 0 (0%) |
| T6 (Jun) | 21.55 | 20.98 | 22.13 | 38 (24%) | 122 (76%) | 0 (0%) |
| Overall Mean | 21.63 | 21.22 | 22.05 | 205 (23%) | 685 (76%) | 6 (0.7%) |

The findings suggest that students' levels of test hope, as measured by the mean test hope Likert scores, fluctuated over the school year. From September to November, levels of test hope initially increased before reaching their highest point in March. However, levels of test hope declined in January, which coincided with the mock examinations, and continued to decrease until June. Despite these fluctuations, the mean test hope Likert scores consistently fell within the average range across the school year, as indicated by the reported scores ranging from MT1=21.44 to MT6=21.55. None of the participants reported no level of test hope. Further analysis and interpretation of these findings in conjunction with other variables and factors can provide valuable insights into students' test-related emotions and their implications for academic performance and well-being.

In the norm group sample studied by Pekrun et al. (2004), the published mean for test hope was found to be M=25.91 (SD=4.93). The levels of test hope observed in this study are within 1 standard

deviation of this published mean. This suggests that while the participants in the Pekrun et al. (2004) study displayed slightly higher levels of hopefulness towards testing compared to the young people in this study, these differences fall within the expected range.

The range analysis presented in Table 6 shows that only six individuals reported levels of test hope that fell within the higher range across the school year. The percentage of individuals reporting lower levels of test hope fluctuated in a wave-like pattern over time, with some individuals shifting from lower levels to average levels of test hope.

The results of the repeated measures factorial ANOVA indicate that there were significant differences in the mean test hope ratings across the six time periods, as indicated by a significant main effect of time, $F(4.113, 357.861) = 3.338$, $p < 0.05$, $\eta 2 = 0.037$. This suggests that test hope levels varied across the school year, with fluctuations observed over time. However, there were no significant interaction effects between levels of reported test hope and school setting, indicating that the school setting did not significantly impact the changes in test hope levels observed over time. This suggests that the differences in test hope levels were not influenced by the type of school setting in which the students were studying.

**Test Pride**

Test pride is the feeling of satisfaction and accomplishment that comes from performing well on a test or exam. It is a sense of pride in one's own abilities and hard work, and the recognition that these efforts have paid off in the form of a high score or grade.

***Table 7: Mean Ratings – Test Pride***

| | | | | Range Analysis | | |
|---|---|---|---|---|---|---|
| | Overall Mean | School A Mean | School B Mean | Low Range *Number (%)* | Average Range *Number (%)* | High Range *Number (%)* |
| T1 (Sept) | 27.55 (Ave) | 26.04 | 29.07 | 31 (18%) | 137 (82%) | 0 (0%) |
| T2 (Nov) | 28.38 (Ave) | 27.50 | 29.27 | 16 (11%) | 127 (89%) | 0 (0%) |

| | | | | | | |
|---|---|---|---|---|---|---|
| T3 (Jan) | 28.46 (Ave) | 27.84 | 29.08 | 10 (7%) | 126 (93%) | 0 (0%) |
| T4 (Mar) | 29.01 (Ave) | 28.78 | 29.25 | 8 (6%) | 127 (94%) | 0 (0%) |
| T5 (May) | 27.75 (Ave) | 26.28 | 29.22 | 27 (18%) | 127 (82%) | 0 (0%) |
| T6 (Jun) | 28.06 (Ave) | 27.07 | 29.06 | 25 (16%) | 135 (84%) | 0 (0%) |
| Overall Mean | 28.20 | 27.25 | 29.16 | 117 (13%) | 779 (87%) | 0 (0%) |

Means at each time point were found to consistently fall within the average range across the school year. From September ($M_{T1}$=27.55) to March ($M_{T4}$=29.01) reported levels of test pride consistently increased before falling, prior the beginning of the examination season in May ($M_{T5}$=27.75). Levels of reported test pride recovered somewhat after the examination season by June ($M_{T6}$=28.06). None of the participants reported no level of test pride.

The published mean for test hope is reported by Pekrun et al. (2004) as being M=31.32 (SD=6.48) for their norm group sample. The overall levels of test pride found in this study are within 1SD of the published mean, provides an indication that although participants in the Pekrun et al. (2004) study we slightly more pride around testing than the young people in this study, these levels are within the expected range.

The range analysis presented in Table 7 indicates that across the school year, no individuals reported levels of test pride that fell within the high range. At T1 (September), 82% (*n*=137)) of participants reported scores that fell within the average range. By T4 (March) this has increased to 94% (*n*=127), with the remaining 6% (*n*=8) reporting low levels of test pride, indicating that over time, individuals who previously reported low levels shifted in their responses to indicate average levels test pride.

A repeated measures factorial ANOVA indicated there was a statistically significant difference in the test pride mean ratings across the six time periods, $F(3.754, 326.604) = 3.266$, $p<0.05$, $\eta^2=0.036$. there was a statistically significant interaction effect between reported test pride levels and type of school, suggesting that students at School B reported lower levels of test pride compared to their counterparts in School B.

The effect size (η2) values of 0.036 and 0.032 for the main effect of time periods and the interaction effect with type of school, respectively, suggest a small effect size. This means that the time periods and type of school accounted for a small proportion of the variance in the test pride scores.

### *Test Relief*

Test relief is the feelings of relief due to the stress and pressure that one may have experienced while preparing for and taking a test. Test relief can be a powerful and positive emotion, as it allows individuals to let go of any tension and anxiety, they may have been holding onto during the testing process. It can also be a form of self-validation and a recognition of one's hard work and effort put into studying for the test.

***Table 8: Mean Ratings – Test Relief***

| | | | | Range Analysis | | |
|---|---|---|---|---|---|---|
| | Overall Mean | School A Mean | School B Mean | Low Range *Number (%)* | Average Range *Number (%)* | High Range *Number (%)* |
| T1 (Sept) | 21.70 (Ave) | 21.72 | 21.67 | 0 (0%) | 168 (100%) | 0 (0%) |
| T2 (Nov) | 21.93 (Ave) | 22.07 | 21.79 | 0 (0%) | 139 (97%) | 4 (3%) |
| T3 (Jan) | 13.56 (Low) | 13.83 | 13.29 | 107 (79%) | 29 (21%) | 0 (0%) |
| T4 (Mar) | 20.94 (Ave) | 20.79 | 21.10 | 0 (0%) | 135 (100%) | 0 (0%) |
| T5 (May) | 17.78 (Low) | 15.72 | 19.84 | 53 (39%) | 84 (61%) | 0 (0%) |
| T6 (Jun) | 20.89 (Ave) | 21.14 | 20.63 | 10 (6%) | 150 (94%) | 0 (0%) |
| Overall Mean | 19.47 (Ave) | 19.21 | 19.72 | 117 (13%) | 779 (87%) | 0 (0%) |

Mean test relief scores remained within the average range throughout the school year, except for two time points in January ($M_{T3}$=13.56) and May ($M_{T5}$=17.78), which coincided with examination periods. During these examination periods, the mean test relief scores were lower, indicating that students may have experienced higher levels of stress and pressure associated with exams. However, after the examinations, the levels of reported test relief quickly recovered to near previous levels,

suggesting that the students were able to overcome the stress and experience relief after completing the exams. None of the participants reported no level of test relief.

The published mean for test hope is reported by Pekrun et al. (2004) as being M=21.59 (SD=4.00) for their norm group sample. Except for the January time point (close to the mock examinations) ($M_{T3}$=13.56), levels of test relief found in this study are within 1SD of the published mean. This provides an indication that although participants in the Pekrun et al. (2004) study reported more relief around testing than the young people in this study, these levels are within the expected range.

The results of the repeated measures factorial ANOVA suggest that there were statistically significant differences in the mean ratings of test relief across the six time periods ($F(4.094, 356.162) = 164.705$, $p<0.05$, $\eta2=0.654$). This indicates that test relief levels varied significantly over time. Additionally, there was a significant interaction effect between test relief and type of school ($F(4.094, 356.162) = 12.513$, $p<0.05$, $\eta2=0.126$), indicating that the school setting had an effect on test relief levels.

Further analysis revealed that students at School A reported lower levels of test relief compared to their counterparts in School B. This suggests that the school environment may have an impact on students' perception of test relief, with some students potentially experiencing less relief compared to students at the other school. The effect size ($\eta2=0.126$) indicates that the type of school accounted for approximately 12.6% of the variance in test relief levels.

### *Test Anger*

Test anger refers to the emotional response of frustration, irritation, and rage that a person experiences before, during and after taking a test or examination. This emotional response can be triggered by various factors, such as feeling unprepared for the test, encountering difficult questions, experiencing time pressure, or feeling like the test is unfair or biased.

***Table 9: Mean Ratings – Test Anger***

| | | | | Range Analysis | | |
|---|---|---|---|---|---|---|
| | Overall Mean | School A Mean | School B Mean | Low Range *Number (%)* | Average Range *Number (%)* | High Range *Number (%)* |
| T1 (Sept) | 23.07 (Low) | 23.20 | 22.95 | 128 (76%) | 40 (24%) | 0 (0%) |
| T2 (Nov) | 23.36 (Low) | 23.49 | 23.23 | 98 (69%) | 45 (31%) | 0 (0%) |
| T3 (Jan) | 24.59 (Low) | 24.77 | 24.40 | 67 (49%) | 69 (51%) | 0 (0%) |
| T4 (Mar) | 23.87 (Low) | 24.00 | 23.74 | 80 (59%) | 55 (41%) | 0 (0%) |
| T5 (May) | 24.76 (Low) | 25.72 | 23.81 | 76 (49%) | 78 (51%) | 0 (0%) |
| T6 (Jun) | 21.58 (Low) | 21.27 | 21.90 | 132 (83%) | 28 (17%) | 0 (0%) |
| Overall Mean | 23.54 (Low) | 23.74 | 23.34 | 581 (65%) | 315 (35%) | 0 (0%) |

The results of the repeated measures factorial ANOVA showed statistically significant differences in the mean test anger ratings across the six time periods, with $F(3.949, 343.575) = 17.198$, $p<0.05$, $\eta 2=0.165$. Specifically, higher levels of test anger were observed in January (MT3=24.59) and May (MT5=24.76), which were the two time points coinciding with examinations. This suggests that students experienced increased levels of test anger during these high stakes testing periods. However, there was no significant interaction effect between test anger and the type of school setting, indicating that the levels of test anger were similar School A and School B. None of the participants reported no level of test anger. It is noteworthy that the mean test anger ratings consistently fell within the low range across the school year, indicating that test anger was generally not a prominent emotion among the students in this study, except during the examination periods in January and May.

Based on the findings from Pekrun et al.'s (2004) study, it appears that participants in the current study reported slightly higher levels of test anger compared to the norm group sample. However, these levels are still within the expected range, as they fall within 1 standard deviation (SD) of the published mean.

The range analysis presented in Table 9 indicates that at the beginning of the school year (T1), 24%

of participants reported average levels of test anger, while the majority (76%) reported low levels. However, by T3 (January), the percentage of participants reporting average levels of test anger increased to 51%, with the remaining 49% reporting low levels. This suggests that over time, individuals who initially reported low levels of test anger shifted in their responses to indicate higher, but still average, levels of test anger. No individuals in the study reported levels of test anger that fell within the high range throughout the school year, indicating that overall levels of test anger remained within the expected range.

Overall, these findings suggest that while participants in the current study reported slightly higher levels of test anger compared to the norm group sample, these levels are still within the expected range and show a trend of increasing test anger levels over time, particularly among those who initially reported low levels of test anger.

### *Test Anxiety*

Test anxiety is a type of performance anxiety that occurs when a person is preparing for or taking a test or exam. It is characterised by feelings of worry, nervousness, fear, and stress that can interfere with a person's ability to perform well on the test. Test anxiety can manifest in physical symptoms such as sweating, rapid heartbeat, stomach upset, and headaches.

***Table 10: Mean Ratings – Test Anxiety***

| | | | | Range Analysis | | |
|---|---|---|---|---|---|---|
| | Overall Mean | School A Mean | School B Mean | Low Range *Number (%)* | Average Range *Number (%)* | High Range *Number (%)* |
| T1 (Sept) | 40.67 (Ave) | 39.32 | 42.03 | 128 (76%) | 40 (24%) | 0 (0%) |
| T2 (Nov) | 41.93 (Ave) | 41.26 | 42.60 | 0 (0%) | 91 (64%) | 52 (36%) |
| T3 (Jan) | 40.60 (Ave) | 39.32 | 41.88 | 3 (2%) | 94 (69%) | 39 (29%) |
| T4 (Mar) | 47.97 (High) | 47.43 | 48.52 | 0 (0%) | 15 (11%) | 120 (89%) |
| T5 (May) | 47.99 (High) | 48.66 | 47.34 | 0 (0%) | 20 (12%) | 143 (88%) |
| T6 (Jun) | 41.91 (Ave) | 41.03 | 42.79 | 0 (0%) | 107 (67%) | 53 (33%) |
| Overall Mean | 43.51 (High) | 42.03 | 44.19 | 131 (14%) | 367 (41%) | 407 (45%) |

The mean test anxiety levels observed in the study showed fluctuations over time. In September (MT1), the mean test anxiety level was 40.67, which was within the average range. In January (MT3), the mean test anxiety level decreased slightly to 40.60, remaining within the average range. However, by March (MT4), after the mock examinations, the mean test anxiety level increased significantly to 47.97, and this elevated level continued up to the commencement of the examination season in May (MT5), with a mean test anxiety level of 47.99. Following the examinations, the reported test anxiety levels dropped back to within the average range. None of the participants reported no level of test anxiety.

Table 8 shows a decrease in test relief scores after the mock examinations in March, indicating higher levels of test anxiety during that period. The scores may then show a further decrease after the actual examinations in May, suggesting a reduction in test anxiety levels post-examination.

Overall, the findings indicate that test anxiety levels varied across different time points in the study, with higher levels observed during the mock examinations and the commencement of the examination season, and lower levels observed before and after these periods, falling back within the average range.

The results of the repeated measures factorial ANOVA indicated statistically significant differences in mean test anxiety levels across the six time periods ($F(3.634, 316.159) = 58.815$, $p<0.05$, $\eta 2=0.403$). This suggests that test anxiety levels changed significantly over time during the study.

There was also a significant interaction effect between reported test anxiety levels and the type of school ($F(3.634, 316.159) = 2.547$, $p<0.05$, $\eta 2=0.028$), indicating that there were differences in test anxiety levels between participants in School A and School B. Specifically, School B participants reported higher levels of test anxiety compared to their counterparts in School A.

When comparing the mean test anxiety levels reported in this study to those reported by Pekrun et al.

(2004), it was found that the participants in this study had slightly lower levels of test anxiety (M=43.51) compared to Pekrun et al. (2004) sample (M=45.54), but this difference was within 1 standard deviation below the mean, indicating that the levels of test anxiety in this study were within the average range.

The range analysis presented in Table 10 indicates that at the beginning of the school year (T1, September), the majority of participants (76%) reported low levels of test anxiety. However, by T4 (March), a large proportion (89%) of participants reported high levels of test anxiety. This pattern is visually illustrated in Figure 13, which shows a quick change in test anxiety levels over time, with a significant increase observed by March.

Overall, these findings suggest that test anxiety levels fluctuated over the course of the study, with a significant increase observed by March, and that there were differences in test anxiety levels between participants in School A and School B. However, the levels of test anxiety in this study were within the expected range compared to previous research by Pekrun et al. (2004).

### *Test Shame*

Test shame describes those feelings of embarrassment, guilt, or inadequacy that a person may experience as they prepare for, take, and reflect on tests or examinations. These feeling may arise from a fear of judgment or criticism from others, or from a sense of personal failure or disappointment.

***Table 11: Mean Ratings – Test Shame***

| | | | | Range Analysis | | |
|---|---|---|---|---|---|---|
| | Overall Mean | School A Mean | School B Mean | Low Range *Number (%)* | Average Range *Number (%)* | High Range *Number (%)* |
| T1 (Sept) | 21.47 | 21.17 | 21.78 | 150 (89%) | 18 (11%) | 0 (0%) |
| T2 (Nov) | 20.62 | 20.66 | 20.59 | 143 (100%) | 0 (0%) | 0 (0%) |
| T3 (Jan) | 23.85 | 27.84 | 19.85 | 76 (56%) | 60 (44%) | 0 (0%) |

| | | | | | | |
|---|---|---|---|---|---|---|
| T4 (Mar) | 21.56 | 21.04 | 22.08 | 120 (89%) | 15 (11%) | 0 (0%) |
| T5 (May) | 28.87 | 29.00 | 28.73 | 15 (10%) | 139 (90%) | 0 (0%) |
| T6 (Jun) | 24.29 | 24.67 | 23.91 | 88 (55%) | 72 (45%) | 0 (0%) |
| Overall Mean | 23.44 | 24.06 | 22.82 | 592 (66%) | 304 (33%) | 0 (0%) |

A repeated measures analysis of variance (ANOVA) revealed that there were statistically significant differences in the mean reported test shame levels across the six time points, $F(5, 435) = 10.231$, $p<0.05$, $\eta2=0.105$. Post-hoc analyses indicated that test shame levels were consistently low throughout the school year, except for May (MT5=28.87) just before the beginning of the examination season. Following the examinations, test shame levels dropped back to within the low range. None of the participants reported no level of test shame. The findings suggest that test shame levels tend to be low among the participants throughout the school year but may increase as the examination season approaches and then decrease after the examinations are completed.

The results of a repeated measures factorial ANOVA indicated statistically significant differences in the mean test shame ratings across the six time periods, $F(3.945, 343.228) = 99.712$, $p<0.05$, $\eta2=0.534$. Post-hoc analyses revealed that initially, 89% of participants (n=150) reported low levels of test shame at the start of the school year in September (T1). This increased to 100% of participants (n=143) reporting low levels of test shame in November (T2). However, across the remainder of the year, test shame levels fluctuated, with 90% of participants (n=139) reporting average levels of test shame in May (T5), just before the beginning of the examination season.

Furthermore, there was a significant interaction effect between test shame and type of school, $F(3.945, 343.228) = 30.997$, $p<0.05$, $\eta2=0.263$. Participants at School A reported higher levels of test shame than their peers in School B. This suggests that the type of school may have an impact on test shame levels, with School A participants reporting higher levels compared to their School B counterparts.

It is worth noting that the overall levels of test shame (M=23.44) reported by the participants in this study were slightly lower than the mean reported by Pekrun et al. (2004) (M=21.92) in their study, but still within the average range. This indicates that the participants in this study may have slightly lower levels of test shame compared to the young people in Pekrun et al. (2004) sample.

### *Test Hopelessness*

Test hopelessness refers to a state of despair, helplessness, or discouragement that individuals may experience when facing a test or exam. This feeling can arise from various factors, including being overwhelmed by the testing process, struggling to understand the subject matter, lacking self-confidence or self-esteem, and encountering repeated failures in past tests. Test hopelessness can lead to a sense of resignation, ultimately causing individuals to lose motivation to adequately prepare for the test or even consider not taking it at all.

***Table 12: Mean Ratings – Test Hopelessness***

| | | | | Range Analysis | | |
|---|---|---|---|---|---|---|
| | Overall Mean | School A Mean | School B Mean | Low Range *Number (%)* | Average Range *Number (%)* | High Range *Number (%)* |
| T1 (Sept) | 24.93 | 25.40 | 24.47 | 127 (76%) | 41(24%) | 0 (0%) |
| T2 (Nov) | 24.46 | 24.78 | 24.15 | 128 (90%) | 15 (10%) | 0 (0%) |
| T3 (Jan) | 27.05 | 27.22 | 26.87 | 56 (41%) | 80 (59%) | 0 (0%) |
| T4 (Mar) | 28.10 | 29.16 | 27.04 | 51 (38%) | 84 (62%) | 0 (0%) |
| T5 (May) | 30.25 | 30.10 | 30.40 | 17 (11%) | 137 (89%) | 0 (0%) |
| T6 (Jun) | 25.69 | 28.09 | 23.30 | 93 (58%) | 67 (42%) | 0 (0%) |
| Overall Mean | 26.75 | 27.46 | 26.04 | 472 (53%) | 424 (47%) | 0 (0%) |

The findings indicate that at the beginning of the school year, up to November (MT2=24.46), levels of reported test hopelessness were within the low range. However, from January (MT3=27.05), levels of test hopelessness increased and reached the average range, with a peak in May (MT5=30.25). None of the participants reported no level of test hopelessness.

It is worth noting that the increase in test hopelessness levels from January to May suggests that as the school year progressed and approached the examination season in May, participants reported higher levels of test hopelessness. This may indicate that the pressure and stress associated with exams may have contributed to an increase in test hopelessness levels among the participants.

The patterns observed in the data suggest that test hopelessness levels tend to be low at the beginning of the school year, increase towards the average range during the year, and reach a peak around the time of the examinations. These findings provide insights into the dynamics of test hopelessness levels across the school year, which can contribute to a better understanding of students' emotional experiences related to exams and assessments.

The results of the repeated measures factorial ANOVA indicate that there were statistically significant differences in the mean ratings of test hopelessness across the six periods, with $F(3.913, 340.390) = 48.251$, $p<0.05$, $\eta 2=0.357$. This suggests that test hopelessness levels varied significantly throughout the school year. Furthermore, there was a significant interaction effect between test hopelessness levels (THL) and type of school, with $F(3.913, 340.390) = 8.530$, $p<0.05$, $\eta 2=0.089$. This indicates that participants at School A reported higher levels of test hopelessness compared to their peers at School B.

It is interesting to note that the overall levels of test hopelessness ($M=26.75$) reported by the participants in this study, as compared to Pekrun et al. (2004) sample participants ($M=22.12$), were slightly lower, falling within the average range. This suggests that the participants in this study may have had relatively lower levels of test hopelessness compared to the young people in Pekrun et al. (2004) study.

The range analysis presented in Table 12 indicates that at the beginning of the school year (T1 in September), 76% of participants (n=127) reported low levels of test hopelessness, which increased to 90% of participants (n=90) reporting low levels in November (T2). However, across the remainder of

the year, test hopelessness levels decreased, with only 11% of participants (n=17) reporting average levels of test hopelessness in May (T5).

These findings suggest that test hopelessness levels tend to be relatively low at the beginning of the school year, increase during the year, and then decrease towards the end of the school year, particularly after the examinations. The significant interaction effect between test hopelessness levels and type of school also highlights the potential influence of school environment on students' experiences of test hopelessness, with School A participants reporting higher levels compared to their peers at School B.

**Discussion**

This research explores how young people's reported test emotion levels changed across the academic year and how these levels differed between two schools located within differing settings. Our findings demonstrate that some test emotion levels, such as Test Enjoyment, Test Hope, Test Pride and Test Anger, remain relatively static across the school year. Others, including Test Relief, Test Anxiety, Test Shame and Test Hopelessness, fluctuate in apparent response to significant testing events such as the proximity to mock and final examinations. Except for Test Hope, all the positive test emotion levels increased following the final examinations, while the negative emotions were found to have decreased. This observed pattern is supported by Mccaldin's (2019) findings, where she reported that examination stress greatest prior to the commencement of the annual exam season and decreased rapidly once it had begun.

As discussed in the findings section, there were no published cut off points for the Test Related Emotions Scales. Table 2 (above) has presented a rubric that was developed by the authors to allow for consistency in describing the changes in test emotions across time. This rubric was aligned with the five-point Likert scale used in the instrument, and it included specific cut-off points that served as a framework for interpreting the results and comprehending the reported changes in test emotions

by the participants.

Young people across the two schools reported differing patterns of test emotions. At School A, young people reported:

(1) Similar levels of Test Enjoyment, Tests Hope and Test Anger,

(2) Lower levels of Test Pride, Test Relief and Test Anxiety, and

(3) Higher levels of Test Shame and Test Hopelessness in comparison to their School B peers.

Accordingly, School B participants reported (i) lower levels of Test Relief, Test Shame and Test Hopelessness, and (ii) higher levels of Test Pride and Test Anxiety.

Previous research has suggested that test emotions, such as test anxiety, are common within school aged populations. Von der Embse, Barterian, and Segool (2013) estimated that test anxiety affects up to 40% of students, whereas Putwain and Daly (2013) report as many as 16% of Year 11 students may experience high levels of test anxiety. Our findings indicate significant fluctuations in reported test anxiety levels across the academic year, ranging from 0% of young people reporting high test anxiety levels in September to 89% reporting such levels in March. It is argued that the varying reported prevalence rates in previous literature may be explained, in some part at least, by factors such as the school setting and the time of year given that the observable pattern suggests that this is linked to testing events like mock exams and the approach of final examinations. Consequently, researchers who are investigating test anxiety and broader test emotions may wish to consider the following factors when conducting their research:

1. **The type of school being studied:** It was found that young people in School A experienced lower levels of test hope, pride, relief and anxiety and higher levels of shame and hopelessness than their School B counterparts. Correspondingly, young people in School B appeared to experience higher levels of anxiety and pride around testing and

similar levels of test enjoyment, hope, and anger to their School A peers.

It is felt that this is an important point as test anxiety is frequently reported in literature as being an inherent, trait factor within the individual, commonly known as a "*within child*" factor (Putwain & Daley,2014). However, these results begin to point to the educational environment itself possibly playing a role in influencing the emotional well-being of children and young people. The researcher also points out various broader factors within the UK educational system that can contribute to the pressures faced by children and young people. These factors include the increased use of testing, accountability systems, changes in course assessment requirements, and more demanding regulatory frameworks (Mansell, 2007).

**2. The time of year the research is conducted:** Test emotion levels such as anxiety, relief, shame, and hopelessness may fluctuate based on the academic calendar and proximity to significant testing events, therefore research findings may vary depending on the time they are conducted.

By carefully considering these factors, researchers can better account for their potential effects on measured test anxiety/ emotion levels, thus drawing more accurate conclusions and making more informed recommendations for addressing the emotional experience of testing in educational settings.

**Implications for practice**

Given the findings that test emotions such as Test Relief, Test Anxiety, Test Shame and Test Hopelessness, appear to vary across time, presenting differently in the schools that participated within this study, professionals working with young people, including teachers, pastoral staff, and educational psychologists, may find it useful to recognise that test emotions may manifest differently

in different individuals and situations.

It is argued that these findings have several implications for practice, including a need to:

1. **Support students' emotional preparedness for test experiences:** Considering the fluctuation of positive and negative test emotions throughout the academic year, it would be beneficial for schools to consider how students can be effectively prepared for the emotional experiences they may encounter. This could involve discussing with students the range of emotions they might experience and how these emotions can vary at different points throughout the year.

   School pastoral staff should be attentive to heightened negative emotions, such as Test Anger, Test Hopelessness, and Test Shame, during crucial periods of the academic year. It may be beneficial for them to devise strategies to support students experiencing fluctuating emotions proactively. This proactive approach aims to minimise the detrimental impact of these emotions on students' academic performance and overall well-being.

2. **Incorporate fluctuating emotional patterns in test anxiety interventions:** As this is the first longitudinal study of test emotions, authors of existing test anxiety and broader emotional literacy and support interventions, such as STEPS (Putwain et al., 2014) BEAT (Kent Educational Psychology Service, 2016) may wish to review their programmes to include help and support for students to understand and recognise that they may experience a range of fluctuating emotional patterns across the school year.

   These findings present a fresh perspective on students test emotions, and it is suggested that reflecting on these perspectives aligns with the SEND Code of Practice and the Health & Care Professions Council (HCPC) standards of proficiency for educational psychologists (HCPC, 2017) as these guidelines emphasise the significance of incorporating the voices of

young individuals in shaping practices. By aligning interventions with the contemporaneous experiences of young people, professionals can ensure that their interventions are more effective and relevant.

3. **Empower educators to foster student well-being and academic performance:** This study has offered significant insights into the dynamic nature of students' test-related emotions throughout the school year. School leaders and teachers could apply similar methodologies to examine the range and intensity of test emotions within their educational environments. The findings presented in this study offers schools a comparator against which they can compare and evaluate the test emotions observed in their specific settings.

By proactively assessing and addressing test emotions, educators can empower students to navigate the emotional landscape associated with tests and exams more effectively. This can potentially enhance students' well-being, motivation, and academic performance.

**Implications for future research**

The context of this research provides fruitful opportunities for future research including examination of possible sociological factors within test anxiety including socio-economic status, parental expectations and social comparison and examination of possible interactions between different emotions. Most useful next steps include:

1. **Examination of individual and group differences in test emotions:** This research has identified broad patterns in test emotions. As such, it is the researchers view that there is a need to further explore individual and group differences in test emotions. This could include:
    a. Identifying the test emotional responses for children and young people with special educational needs or who have been identified as being gifted and talented.

**b.** Investigating possible differences brought about through different test or examination models such as oral language assessments, practical examination, and modular courses. This will offer a more nuanced understanding of young people's experiences that can inform tailored interventions to address test anxiety effectively.

**c.** Exploring how test emotions vary according to the subjects being studied. It is speculated, based on the findings of Mccaldin (2020), that it is possible that certain subject areas, for example, mathematics, physics, history, geography, and art, may induce differing patterns of test emotions.

2. **Timing of interventions:** Given the findings that test emotions can vary across the year, schools may wish to consider when might be the optimal time for delivering test anxiety interventions, considering factors such as testing periods, age or grade level of students, and the specific type of intervention being implemented. Identifying the ideal timing for interventions will support schools planning and possibly enhance their effectiveness in reducing test anxiety.

*3.* **School culture:** The culture of a school, comprising its values, norms, and expectations, can influence the emotional state of students regarding tests (Berman & McLaughlin, 1975; Turner et al, 2002). For instance, a school that places a strong emphasis on test scores or fear appeals (Mccaldin, 2020) may generate a more stressful environment for students, whereas a school that prioritises student well-being may have a more affirmative impact on students' test anxiety. Understanding the role of school culture in test emotions can assist schools in building supportive environments that foster positive emotions and mitigate stress.

**Limitations**

We have identified four possible limitations to the research reported within this paper. First, as this is the first longitudinal study of test emotions, there is a clear need to examine the replicability of our

findings across a broader range of settings and time points. Possible variations may exist regionally, nationally, or in different educational settings such as independent, international, and special schools.

Second, and linked to the above, the study was conducted in two schools, and the results may not fully represent the national population range. However, the research has characterised the two settings and selected schools with distinct differences to provide a basis for comparison.

Third, the uncertainty surrounding academic assessments and examinations during the COVID-19 pandemic has also increased test anxiety (Schafer et al., 2020; Chen et al., 2021). Changes in the format of examinations have also been shown to exacerbate test anxiety in students who may feel unprepared or unsure about what to expect in these modified testing conditions (Putwain & Aveyard, 2018). It is difficult to predict how these patterns may change over time; however, as schools and examination practices return to established and familiar patterns, it is anticipated that such uncertainties will be reduced with a resulting reduction in test anxiety which may also influence the replicability of the test emotions levels reported in this study.

Fourth, we acknowledge that there is a potential source of variance introduced to the study given that participants may, through their choice of GCSE subject options have experienced a variety of testing and examination situations at different time points that could have included mock examinations, language oral assessment, practical assessments, or modular assessments within BTEC subjects. As such, this introduces an element of heterogeneity into the data as participants may have responded with these different tests in mind. While we do acknowledge this as a possible limitation of this study, we also argue that this broad view is an appropriate first step in researching end-of school students' test emotions over time. Future research may reveal differing emotional responses to different testing and examination regimes, and/ or possible and differing patterns in test emotions between sub-groups such as those with special educational needs or though who are identified as being gifted and talented.

**Disclosure statement**

The authors reported no potential conflict of interest.

**References**


Ajilchi, B., & Nejati, V. (2017). Executive functions in students with depression anxiety, and stress symptoms. *Basic Clinical Neuroscience*, 8(3), 223-232.

Bartolucci, F., Bacci, A., & Gnaldo, M. (2016). Statistical Analysis of Questionnaires: A Unified Approach Based on R and Stata. Net York: Routledge.

Bentley, H., Fellows, A, Glenister, S. & Mussen, N. (2019). How safe are our children?: An overview of data on adolescent abuse. London: NSPCC.

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

Chen, J., Wang, X., Lin, Y., & Li, Y. (2021). The Impact of COVID-19 Pandemic on Children's Learning and Anxiety: An Empirical Study Based on the Experience of China. Journal of Psychoeducational Assessment, 39(3), 279-288. doi: 10.1177/0734282920962938.

Berman, H., & McLaughlin, V. (1975). Predicting test anxiety and achievement from student and classroom variables. Journal of Educational Psychology, 67(6), 717-722.

British Psychological Society. (2014). *BPS Code of Human Research Ethics* (2nd ed.). https://www.bps.org.uk/news-and-policy/bps-code-human-research-ethics-2nd-edition-2014

Cassady, J. C., & Johnson, R. E. (2002). Cognitive test anxiety and academic performance. Contemporary Educational Psychology, 27(2), 270-295.

Damer, D & Melendres, L. (2011) “Tackling Test Anxiety”: A Group for College Students. *The Journal for Specialists in Group Work*, 36:3, 163-177

Erdfelder, E., Faul, F., Lang, A., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175-191.

Hair, J.F., Risher, J.J., Sarstedt, M. and Ringle, C.M. (2019), When to use and how to report the results of PLS-SEM, *European Business Review*, Vol. 31 No. 1, pp. 2-24.

Hembree, R. (1988). Correlates, causes, effects, and treatment of test anxiety. Review of Educational Research, 58(1), 47-77.

Health and Care Professions Council. (2017). Standards of proficiency - Educational psychologists. Retrieved from https://www.hcpcuk.org/globalassets/resources/standards/standards-of-proficiency---educational-psychologists.pdf

Hsieh, Y. P., & Lee, J. C. (2017). Influence of parenting styles on test anxiety: The role of perceived academic competence as mediator. Asia Pacific Education Review, 18(4), 585-593.

IBM Corp. (2020). IBM SPSS Statistics for Windows (Version 27.0) [Computer software]. IBM Corp.

Keogh, E., Bond, F. W., & Flaxman, P. E. (2006). Improving academic performance and mental health through a stress management intervention: Outcomes and mediators of change. *Behaviour Research and Therapy*, 44(3), 339-357

Li, C., & Lai, M. (2011). The influence of teacher and peer relationships on students' classroom engagement and everyday functioning. Journal of Youth and Adolescence, 40(2), 164-175.

Mansell, W. (2007). *Education by numbers: The tyranny of testing*. London: Methuen Publishing

May F, Schaffer GE, Allen KA, Berger E, von Hagen A, Hill V, Morris ZA, Prior S, Summers D, Wurf G, Reupert A. Perspectives of practicing school psychologists during COVID-19: A multi-country, mixed methods investigation. Sch Psychol Int.

Mccaldin, T. (2020). The experiences of students preparing for GCSE examinations: An interpretative phenomenological analysis. University of Manchester, United Kingdom.

Miles, M. B., & Huberman, A. M. (1994). Qualitative data analysis: An expanded sourcebook (2nd ed.). Sage: London.

NSPCC (2016) How safe are our children? London: NSPCC

Office for National Statistics. (2019). The English Indices of Deprivation 2019. [IoD2019]. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/835115/IoD2019_Statistical_Release.pdf

Office for National Statistics. (2019).

Patton, M. Q. (2015). Qualitative research & evaluation methods: Integrating theory and practice (4th ed.). Thousand Oaks, CA: Sage.

Pekrun, R., Muis, K.R., Frenzel, A. C., & Goetz, T. (2018). *Emotional at School*. New York: Routledge.

Pekrun, R., Elliot, A. J., & Maier, M. A. (2009). Achievement goals and achievement emotions: Testing a model of their joint relations with academic performance. Journal of Educational Psychology, 101(1), 115-135.

Pekrun, R., Muis, K.R., Frenzel, A. C., & Goetz, T. (2018). Emotions at School. New York: Routledge.

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.

Peleg, O. (2009). Test anxiety, academic achievement, and self- esteem among Arab adolescents with and without learning disabilities. *Learning Disabilities Quarterly*, 32, 11-20.

Putwain, D. W. (2007). Researching the relationship between achievement goals, academic emotions and achievement-related outcomes. Educational Psychology, 27(5), 507-535.

Putwain, D. W., & Aveyard, B. (2018). Is perceived uncertainty a transdiagnostic factor in test anxiety? Journal of Educational Psychology, 110(6), 829–837.

Putwain, D., & Daly, A. L. (2014). Test anxiety prevalence and gender differences in a sample of English secondary school students. *Educational Studies, 40*(5), 554–570.

Rodway, C., Tham, S.G., Ibrahim, S., Turnbull, P., Windfuhr, K., Shaw, J., Kapur, N., Appleby, L., (2016) Suicide in children and young people in England: A consecutive case series. *Lancet Psychiatry*. 3(8), 751-759.

Salend, S. J., & Garrick Duhaney, L. M. (2005). The impact of socialization and academic factors on anxiety among students with disabilities. Journal of School Psychology, 43(4), 309-325.

Schneider, M., & Preckel, F. (2017). Variables associated with achievement in higher education: A systematic review of meta-analyses. Psychological Bulletin, 143(6), 565-600.

Seipp, B. (1991). Anxiety and academic performance: A meta-analysis of findings. *Anxiety Research*, 4(1), 27–41

Sirin, S. R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. Review of Educational Research, 75(3), 417-453.

Stobart, G. (2008) Testing times: The uses and abuses of assessment. New York, Routledge.

Thomas, C. R. and Gadbois, S. A. (2007), Academic self-handicapping: The role of self-concept clarity and students' learning strategies. *British Journal of Educational Psychology*, 77: 101–119.

Turner, J. C., Midgley, C., Meyer, D. K., Gheen, M., Anderman, E. M., Kang, Y., ... & Patrick, H. (2002). The classroom environment and students' reports of avoidance strategies in mathematics: A multimethod study. Journal of Educational Psychology, 94(1), 88-106.

Ursache, A., Raver, C.C., (2014) Trait and state anxiety: relations to executive functioning in an at-risk sample. *Cognition & Emotions*. 28(5):845-55.

Wang, M. T., Eccles, J. S., & Kenny, S. (2013). Not lack of ability but more choice: individual and gender differences in choice of careers in science, technology, engineering, and mathematics. Psychological Science, 24(5), 770-775.

Yin, R.K, (2013). Case Study Research & Applications. Sage: London.

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

# Paper 3: The dissemination of research evidence to professional practice

Word count: 5364

## Introduction

This paper investigates the knowledge transfer process through research dissemination and its interrelation with evidence-based practice and practice-based evidence. It also examines possible research impact and how this might be evaluated. The paper is structured into two main sections, with Part 1 elucidating the theoretical foundation of knowledge transfer. In contrast, the Part 2 investigates the practical strategies for disseminating the research outcomes presented in Papers 1 and 2. Specifically, the paper examines dissemination tactics at various levels, including the whole school, organisational, and professional levels. It proposes a detailed approach for promoting and evaluating the dissemination and impact of the research within this thesis.

### *The researcher's role in the transfer of knowledge*

In the context of educational psychology, the transfer of knowledge around academic research refers to the process of sharing research findings, insights, and expertise with clients, peers, parents, schools, local authorities, and broader society (Melloe & Stoloff, 2017). It involves communicating research in accessible and understandable ways and ensuring that research is relevant and responsive to the needs of stakeholders (Taylor, 2019). Transfer of knowledge is a critical component of academic research; indeed, Alberts & Godlee (2017) argue that the dissemination of research findings is not only an ethical obligation but also a necessary step in the scientific process that helps ensure the reliability and validity of the research as well as highlighting the practical use of research to society.

Linked to the above, Kelly (2022) identifies three underpinning reasons behind research professionals' duty to engage in knowledge transfer which can be summarised as being:

(1) **The obligation to public/private research commissioners:** As recipients of public and private funding, research professionals have an obligation to ensure that their work has a

positive impact on society. Knowledge transfer is a key mechanism for achieving this impact, as the process around which researchers share their findings with practitioners, policymakers, and other stakeholders who can apply this knowledge to real-world contexts.

(2) **Dissemination as an essential part of the research process:** Knowledge transfer is an essential component of the research process itself and as such research professionals are responsible for not only producing knowledge but also disseminating and communicating it effectively. This requires researchers to engage with stakeholders, communicate their findings in accessible and understandable ways, and provide opportunities for feedback and discussion.

(3) **Ensuring that research findings are relevant and responsive to the needs of individuals, organisations, and society:** By engaging in knowledge transfer, research professionals can gain insights into the challenges faced by practitioners and other stakeholders, and tailor their research to address these challenges. This can help to ensure that research findings are more likely to be adopted and implemented in practice, leading to positive outcomes for individuals and society.

Research training is an important component of doctoral educational psychology programmes, which aim to develop trainee educational psychologists'(TEPs') research knowledge and skills and to support their contribution to the advancement of knowledge in the field through publication. This training helps to prepare TEPs for careers as active researchers and practitioners, this being what Corrie & Lane (2006) and Fallon, Woods & Rooney (2010) refer to as a scientist-practitioner role.

### The Importance of Evidence-Based Practice and Practice-Based Evidence

As effective knowledge transfer involves the dissemination of research findings to practitioners and the translation of these findings into practice, this links closely to Evidence-Based Practice (EBP) and Practice-Based Evidence (PBE). These two interconnected approaches aim to improve client

outcomes through the integration of scientific knowledge and research evidence with real-world practice (Lane & Corrie, 2006).

### *Evidence-Based Practice*

The American Psychological Association (APA) defines EBP as "the integration of the best available research with clinical expertise in the context of patient characteristics, culture, and preferences" (APA, 2006). Evidence-Based Practice (EBP) originates from Cochran's (1979) work on randomised control trials within the medical field. Cochran (ibid) argued that clinical decisions should be based on the best available evidence rather than solely on practitioner experience or intuition. Over time, EBP has gained widespread acceptance within various fields, including education, social work, and psychology (Lane & Corrie, 2006). Today, psychology practitioners are encouraged to review and use methods and approaches that have been systematically researched and demonstrate clear and effective outcomes (Dunsmuir et al., 2009). Indeed, the Health and Care Professions Council (HCPC) has embedded EBP as an expected standard of proficiency for practitioner psychologists, requiring them to engage in evidence-based and evidence-informed practice and systematically evaluate their practice (HCPC, 2016, standard 12.1).

EBP is underpinned by the hierarchy of evidence proposed Scott, Shaw & Joughin (2001). This hierarchy, ranges from systematic reviews of randomised controlled trials to individual opinions and informs and supports professional practice by providing an evaluative framework that identifies the most reliable and valid research approaches (Frederickson, 2002). The hierarchy emphasises the importance of rigorous research methods, placing the most reliable and valid approaches to research at the top. As a result, the following order is proposed:

(1) Several systematic reviews of randomised controlled trials.
(2) Systematic review of randomised controlled trials.

(3) Randomised controlled trials.

(4) Quasi-experimental trials.

(5) Case control and cohort studies.

(6) Expert consensus opinion.

(7) Individual opinion.

Scott, Shaw & Joughin (2001) argue that this hierarchy reflects the scientific approach and recognises rigorous research as good quality evidence. Indeed, Fox (2003) describes randomised controlled trials as the '*gold standard'* in research evidence because they generally involve multiple participants randomly allocated to control or experimental groups that explore the efficacy of interventions (Frederickson, 2002). Accordingly, the hierarchy states that qualitative research, such as case studies or expert consensus and personal views, should be considered a less sophisticated form of less reliable evidence – a view contested by qualitative researchers such as Flick (2014) and Bazeley (2021).

***Practice-Based Evidence***

In contrast, Practice-Based Evidence (PBE) focuses on generating research evidence from within daily practice settings (Lane & Corrie, 2006). PBE acknowledges the unique context of each practice situation, seeks to capture real-world interventions' complexity, and encourages practitioners to reflect on their experiences and systematically evaluate their work and its contribution to the evidence base (Mcniff & Whitehead, 2011).

The PBE approach recognises that claims to knowledge gained through the experience of practitioners is a valuable source of evidence, and that this knowledge can contribute to the development and refinement of effective psychological practice. By capturing the complexity of real-world practice, PBE provides a more comprehensive understanding of the impact of interventions and contributes to the development of evidence-based practice that is tailored to the needs of specific

populations and settings (Mcniff & Whitehead, 2011).

Barkham, Hardy & Mellow-Clark (2010) highlight the complementary relationship between EBP and PBE as being valuable in developing a comprehensive evidence base within the field of educational psychology. In general, they suggest that while EBP ensures that practitioners use well-established, research-backed methods and interventions, PBE contributes insights and information about the effectiveness of these methods in diverse and complex real-world contexts. At a more detailed level, Mcniff & Whitehead's (2011) meta-analysis identifies four specific benefits to taking a combined PBE/EBP approach to professional practice, these being:

(1) **Improved outcomes for children and young people:** EBP supports educational psychologists in employing interventions and strategies that are based on rigorous research, increasing the likelihood of positive outcomes for individuals and communities. Additionally, PBE offers insights into the effectiveness of these methods in diverse and complex real-world contexts, allowing for the adaptation and refinement of interventions based on actual practice.

(2) **Informed decision-making:** Integrating available research with clinical expertise and practice, EBP allows educational psychologists to make well-informed decisions about their interventions and strategies. PBE offers valuable feedback on the applicability of these methods in various contexts, further informing decision-making processes.

(3) **Enhanced credibility:** EBP and PBE help to support educational psychology's credibility as a discipline by demonstrating that interventions are grounded in research evidence and informed by real-world practice. This can lead to increased trust and collaboration with other professionals and greater acceptance and understanding from clients and stakeholders.

(4) **Continuous learning and development:** The ongoing evaluation and adaptation of interventions through EBP and PBE foster a culture of continuous learning and professional development for educational psychologists. This supports and encourages practitioners to

stay updated on the latest research findings, refine their skills, and provide better client services.

In summary, incorporating EBP and PBE into educational psychology practice helps bridge the gap between research and practice, ensuring that interventions are theoretically grounded and practically relevant. This approach fosters the development of more effective and targeted interventions, ultimately leading to improved outcomes for individuals, families, and educational systems (Godfrey & Brown, 2019).

***Dissemination.***

Smith (2018) highlights that the "*The value of a scientific paper is not in its publication, but in the dissemination of its content*", pointing out that research publication and dissemination are related concepts, but they refer to very different aspects of the research process. Indeed, in this context, research publication refers to sharing research findings with the academic community through peer-reviewed journals, conferences, and reports. In contrast, research dissemination aims to make research findings more widely accessible to practitioners, policymakers, and the public, ensuring that research is relevant and responsive to stakeholders' needs. The former focuses on advancing knowledge in a particular field, while the latter focuses on making research have a positive impact on society.

Several strategies have been identified as being effective in disseminating research findings. These include:

(1) **Targeted communication:** Tailoring the communication of research findings to the specific needs and interests of the intended audience is seen as being crucial for ensuring their understanding and engagement. Depending on the audience, this may involve using different formats, such as policy briefs, infographics, or executive summaries.

(2) **Accessible language:** Using clear, concise, and jargon-free language can help make research findings more accessible and understandable to a broader audience, including practitioners who may not have extensive research backgrounds.

(3) **Collaboration:** Encouraging collaboration between researchers and practitioners can facilitate the exchange of knowledge and expertise, leading to more effective dissemination and application of research findings. This can be achieved through joint conferences, workshops, and other networking opportunities.

(4) **Timely dissemination:** Sharing research findings as soon as they become available can ensure practitioners can access the most up-to-date evidence to inform their practice.

(5) **Utilising multiple channels**: Various channels for dissemination, such as peer-reviewed journals, conference presentations, webinars, social media, and other online platforms, can help reach a broader audience and increase the visibility and impact of research findings.

(6) **Stakeholder involvement:** Engaging relevant stakeholders, such as policymakers, practitioners, and clients, throughout the research process can help ensure that findings are relevant, applicable, and more likely to be utilized in practice.

***Research impact***

Research impact refers to the effect that research has on society, the economy, the environment, and other domains beyond academia. It encompasses the measurable effects of research on policy, practice, and public discourse, and includes both positive and negative outcomes. Understanding the impact of research is important for evaluating research effectiveness and value and can be understood in various dimensions, including:

(1) **Academic impact:** This refers to the contribution of research to advancing knowledge and understanding within a specific field, often assessed through citation metrics, publication counts, and other bibliometric indicators.

(2) **Practical impact:** This relates to the influence of research on professional practice, such as changes in practitioners' behaviours, adoption of evidence-based interventions, or improvements in service delivery.
(3) **Policy impact:** Research can impact policy by informing the development, implementation, and evaluation of policies or by influencing the decision-making processes of policymakers.
(4) **Social impact:** This encompasses the broader effects of research on individuals, communities, and society, such as improvements in well-being, education, or social inclusion.
(5) **Economic impact:** Research can contribute to economic growth and development by driving innovation, creating new industries, or informing more effective resource allocation.

***Research Evaluation***

Research evaluation refers to the process of assessing the quality, relevance, and impact of research. It involves the systematic and objective assessment of research outputs, such as publications, grants, patents, and other forms of knowledge dissemination (Kelly, 2022). Research evaluation is conducted for various purposes, including funding allocation, promotion and tenure decisions, quality assurance, and strategic planning.

Research evaluation methods can include bibliometric analysis, peer review, expert assessment, and stakeholder engagement. The evaluation criteria may vary depending on the specific context and objectives of the evaluation, but may include measures of academic impact, practical impact, policy impact, social impact, and economic impact.

Hoffman (2021) suggests the following are some possible evaluation methods that may be relevant to research dissemination:

(1) **Surveys:** A survey can be conducted to evaluate the impact of the dissemination plan on the target audience. The survey can be distributed to academics, professionals, policymakers, and

the public to gather feedback on the effectiveness of the dissemination plan (Cohen, Manion, and Morrison, 2011).

(2) **Focus groups:** Focus groups can be conducted with educators, school administrators, and pastoral staff to gather feedback on the impact of the research findings on their practice. The focus groups can be conducted at different intervals to evaluate the long-term impact of the research on practice (Flick, 2014).

(3) **Interviews:** Interviews can be conducted with policymakers to gather feedback on the impact of the research on policy decisions. This will enable the evaluation of the impact of the research on policy decisions and enable policymakers to provide feedback on the dissemination plan (Flick, 2014).

Overall, research evaluation is a critical component of research management and governance, as it provides information on the quality and effectiveness of research, and helps to inform decision-making related to funding, career advancement, and strategic planning.

**Conclusion**

The transfer of knowledge in academic research is important for ensuring that research has a positive impact on society. The integration of evidence-based practice (EBP) and practice-based evidence (PBE) approaches into educational psychology practice fosters the development of more effective and targeted interventions, ultimately leading to improved outcomes for individuals, families, and educational systems. Dissemination of research findings is an essential component of effective knowledge transfer, which involves communicating research in accessible and understandable ways and ensuring that research is relevant and responsive to the needs of stakeholders. Several strategies have been identified as effective in disseminating research findings, including targeted communication, accessible language, collaboration, timely dissemination, utilising multiple channels, and stakeholder involvement.

Research impact refers to the effect that research has on society, the economy, the environment, and other domains beyond academia. Understanding the impact of research is crucial for evaluating its effectiveness and value. Research evaluation refers to the process of assessing the quality, relevance, and impact of research. It involves the systematic and objective assessment of research outputs, such as publications, grants, patents, and other forms of knowledge dissemination. Research evaluation is conducted for various purposes, including funding allocation, promotion and tenure decisions, quality assurance, and strategic planning.

The researcher's role in the transfer of knowledge involves not only producing knowledge but also disseminating and communicating it effectively. Researchers have an obligation to ensure that their work has a positive impact on society, and knowledge transfer is a key mechanism for achieving this impact. Educational psychology training programs aim to develop trainee educational psychologists' (TEPs') research knowledge and skills and support their contribution to the advancement of knowledge in the field through publication, preparing them for careers as active researchers and practitioners. Overall, effective knowledge transfer, dissemination, and research evaluation are critical components of academic research, enabling research findings to have a positive impact on society.

## Part 2: A summary of the policy, practice, research development and implications from the research at school, wider organisational, and professional levels.

### Introduction to Part 2.

The second part of this paper briefly reviews the implications for practice from Papers 1 & 2 at the whole school, organisational, and professional practice levels. It then goes on to propose a dissemination and promotion strategy based on Harmsworth et al.'s (2000) dissemination framework. It concludes by outline some of the possible evaluative activities that would be suitable for this work.

***Summary of principal thesis findings.***

This thesis is comprised of Paper 1: A systematic literature review that has identified and critiqued test anxiety measures used worldwide over a twenty-two-year period. Paper 2 reported an empirical study examining Year 11 students' test anxiety and broader test-related emotions across a school year as they prepared for their GCSE examinations.

***Students:***

- That while some test related emotions remained stable over time, others such as test anxiety, test relief, test shame, and test hopelessness, varied in response to important testing events. Examples of such events were the mock examinations held in January at both schools and the final examinations in May/June. These events included the actual tests themselves as well as the subsequent period leading up to receiving the results.

***Students & Teachers*:**

- It also identified that young people in different school settings experience test emotions differently, possibly being linked to school culture and wider context factors.

***Researchers:***

- Issues within the evidence base due to the age of test anxiety measure's standardisation groups.
- Construct validity concerns due to wide variation in language and terms used to describe the various aspects of test anxiety.
- And a lack of incorporation of young people views and perspectives around test anxiety.

***Implications at organisational and professional levels***

The dissemination of this research to educational professionals is important because it demonstrates

that test anxiety and test emotions vary significantly across the school year. As such, adults working in schools can use this evidence to show young people that their feelings, such as test anxiety, test hope and test shame, are common and a normal part of the examination experience. Further, it highlights the need for more international research around gathering children's and young people's voices around test anxiety to ensure that it reflects their contemporary experiences and views. In doing so, it is hoped that colleagues will reflect on their ways of working and refine their practice, expecting that this will positively influence professional practice and improve learners' lives. More specifically, effective dissemination could:

a) Demonstrate that levels of test anxiety and test emotions vary across the school year and that, as such, this helps explain some of the wide variances in prevalence levels reported within test anxiety literature (von der Embse, Barterian & Segool, 2013; Putwain & Daley, 2014).

b) Provide evidence to young people and educational professionals that feelings such as test anxiety, hope, and shame are a normal part of the examination experience.

c) Raise awareness in schools of the deleterious effects of test anxiety and challenge a populist narrative that young people are seen as 'the snowflake generation

d) Challenge schools and broader educational professionals to ensure that the emotional needs of children and young people in relation to test anxiety are being met (given their position, EPs can challenge the status quo and support and encourage change within schools).

e) Through collaborative working through consultations, promote systemic change around recognising test anxiety within schools.

f) Proposing the need for further research around gathering children's and young people's voices around test anxiety to ensure that it reflects their experiences and views.

g) A recognition that more international research is necessary.

***School level***

The literature review findings have relevance for research site staff as it highlights that there are suitable measures for identifying test anxiety within such setting. The continued use of such measures can help to provide staff with an insight into the levels in their setting in relation to those found within in the previous year at any given point in time across Year 11.

As supporting children and young people around test anxiety is viewed as a whole school approach, and one that can be facilitated by individual staff members, it seems important for school staff and senior leaders to measure the effectiveness of any interventions or developments in the way the school has sought to support its young people, any impact on czstudents' wellbeing and any possible wider effects on exam performance.

At the individual level, it was considered important that the participants who took part in empirical data gathering be made aware of the patterns within the data as this could have a positive effect in them being able to see that the emotions that they may have felt, we also felt by their wider peer group and that this was a normal reaction to the stress of school examinations.

For the adults within the setting, it was also felt that the research data highlighted that young people in Year 11 were likely to express emotions at different points within the school year. For example, it might be expected that young people might experience a heightened sense of negative emotions, such as shame and anger, as well as higher levels of physiological feelings in the two months leading up to the examinations. Having this information can help staff empathise with student feelings and help both staff and the school reflect on how they might respond in advance of these expected emotional peaks.

The researcher has chosen to give feedback on findings to participants using a short letter

incorporating an infographic and this will be developed as part of the dissemination process. This dual-feedback method was chosen because it provided young people with the option of how they chose to engage with the research findings. Email feedback from a small number of young people was positive, suggesting that visual feedback was particularly impactful. Whilst it was important from an ethical perspective not to raise expectations of change, the letter was designed to convey that recommendations had been made to relevant persons with the ability to make change happen.

***Organisational level***

The findings of the systematic literature review also highlight implications for practice at the organisational level. It would be useful for educational psychology service staff to know there are measures of test anxiety suitable for supporting students in the school environment. Some of the measures identified in the literature review may also be of interest educationalists within examination bodies and third sector organisations such as charities and organisations that aim to support children and young people.

There is also an implication for educational psychology services to provide planned and targeted training for the psychologists within their service. Regardless of whether the service intends to use or recommend any measures of test anxiety, the findings of Paper 2 will be of interest to psychologists working with Year 11 students as it offers possible explanations around how a young person might be feeling in relation to upcoming examinations. Educational psychologists are in a unique position to inform schools and setting about the wide range of systemic and targeted work they may be able to offer to schools around test anxiety. Sharing successful examples of how test anxiety interventions can be made feasible in a traded service model can perhaps give educational psychologists an opportunity to have a conversation with schools about commissioning this type of work.

Educational psychology services can also support schools in identifying gaps in their approach to supporting students who are experiencing detrimental levels of test anxiety by using an appropriate

test anxiety measure such as the Multidimensional Test Anxiety Scale (MTAS) (Putwain et al., 2020). Educational psychology services could also support schools in identifying resources, strategies and interventions that would enable them to provide better opportunities for students in the build-up to examinations.

### *Professional level*

There are also research implications at a wider, professional level. For instance, the findings of the literature review are likely to be of interest to qualified EPs and school psychologists, both in the UK and internationally. Paper One provides an overview of test anxiety measures used in research and identifies areas that vary in degree between the measures, such as their underpinning psychological constructs along with related validity and reliability data.

Paper 2 investigates and tracks the test emotions of young people across Year 11. The data collected provides an understanding of young people's emotional state in relation to tests, and this provides opportunities for EPs to co-create meaningful and practical ways of supporting the school and setting staff in using psychology that promotes recognition and affirmation of young people's feelings and opportunities for positive change to those students experiencing test anxiety.

From Paper 2, it would also be useful for TEPs who are enrolled on initial professional training courses to be provided with information about test anxiety measures and young people's test emotions. Paper 1 provides an up-to-date overview of available measures which TEPs may wish to add to their service portfolio and/or signpost to schools and colleagues. The recommendations for future research may also be helpful for TEPs who are considering carrying out doctoral research on test anxiety, as several points to future research opportunities that emerged from the work in this area. Examples include the need for research at the international level on young people's experiences and views around test anxiety as well as further research using Paper 2's gathered data set to further explore the possible interactions between test emotion.

## *Table 1: Dissemination Strategy.*

| Objective of Dissemination. | Target Audience | Key Messages | Communications Channels | Timeline |
|---|---|---|---|---|
| **Stage 1: Awareness -** Sharing of research findings with a wide range of audiences. | | | | |
| 1. To share the above research findings to as wide a general audience as possible. Specifically, this should focus on:<br>a) That young people's test emotions can change across the school year.<br>b) Young people may experience increasing levels of test anxiety, test shame and feelings of hopelessness, particularly as mock and final examinations draw closer. | 1. Students<br><br>**Research Finding A.** | • It is normal to feel a heightened sense of anxiety, shame, and hopelessness in the build-up to mock and final examinations.<br>• These feelings can build up over time, but having an awareness of them can help you to understand these feelings.<br>• Identify some strategies to help you reduce your anxiety levels. Develop a plan to revise in a way that works for you. | • Social media – Instagram, Tik Tok – 30 second info bust.<br>• Sharing information in schools:<br>o Develop assembly for Year 11 Students.<br>o Develop pupil handouts.<br>o Develop posters for display in schools. | • November / December 2023 – ahead of the mock examinations.<br>• March / April 2024 – ahead of the GCSE examination season. |
| | 2. Parents<br><br>**Research Finding A.** | • Young people's feelings of test anxiety shame and hopelessness appear to be linked to important examinations.<br>• Focus on ways to support then that can help them to relax, manage their anxiety levels and cope with the feelings of shame and hopelessness associated with examinations.<br>• Suitable supportive strategies might include:<br>o Encouraging open conversation about test anxiety and emphasise that it is a common experience for young people.<br>o Encourage a growth mindset, emphasising that failure and mistakes are opportunities to learn and grow rather than reasons to feel ashamed.<br>o Recognise and praise students' achievements, progress, and effort to build their self-esteem and reduce feelings of shame | • Information leaflets for parents (to be sent home/shared by schools).<br>• PDF version of the leaflets should be made available to be shared electronically via school's website and social media channels.<br>• Articles for MumsNet.<br>• Online petitions. | • September 2023 |
| | 3. School staff – Teachers,<br><br>Teaching Assistants,<br><br>Emotional Literacy Support<br><br>Assistants, Support Staff,<br><br>Supply Staff.<br><br>**Research Findings A & B** | • Schools may wish to consider activities that may reduce negative test emotions and that support the development of positive test emotions and more general well-being.<br>• Schools might consider teaching students various stress reduction techniques, such as deep breathing, progressive muscle relaxation, and mindfulness meditation, to help them manage test-related stress and associated feelings of shame.<br>• Foster a school culture that that avoids contributing to anxiety and stress and values learning and growth over test scores, helping to reduce the pressure and shame associated with exam performance. | • Articles for Teach Secondary Magazine.<br>• Sharing key information vis social media such as Facebook and LinkedIn. | • November / December 2023 – ahead of the mock examinations.<br>• March / April 2024 – ahead of the GCSE examination season. |
| **Objective** | **Target Audience** | **Key Messages** | **Communications Channels** | **Timeline** |
| **Stage 2: Targeting** – A focus on audiences who may benefit from the research. | | | | |
| 1. To share research findings & raise awareness of build-up of negative test emotions.<br>2. To seek and develop professional networks that may support the sharing of good practice around test anxiety in this area. | 1. Educational Psychologists, Assistant EPs, School Psychologists.<br>2. CAMHS Professionals – Clinical Psychologists. | • Year 11 students' feelings around texting can increase levels of anxiety as well as feelings of shame and hopelessness. These appear to build-up over time as mock and final examinations get closer.<br>• Schools should be encouraged to recognise that for some there may be a facilitative effect, whereas for other this can be debilitating.<br>• Support schools to initially provided a whole school level support taking an informing, recognising, understanding and emotional management approach.<br>• Where needed, schools may wish to offer individualised support for students who struggle with high levels of negative test emotions, including | • Publish papers in professional journals.<br>• Present at EP and educational conferences.<br>• Appropriate social media, e.g. EPNet. | • Ongoing form September 2023 |

| | | | | |
|---|---|---|---|---|
| 3. To promote further research in this area. | **Research Findings A, B, C, D & E.** | counselling, mentoring, or academic support to address the root causes of their emotional distress.<br>• Encourage a school culture that values learning and growth over test scores, helping to reduce the pressure and shame associated with exam performance.<br>• Encourage school to actively gather students' views and perspectives around test anxiety.<br>• Communicate that while test anxiety measure is useful as monitoring tools, they may not provide an accurate measure of TA levels given that they may rely on outdated standardisation samples and a wide range of TA language and terms. | | |
| **Objective** | **Target Audience** | **Key Messages** | **Communications Channels** | **Timeline** |
| **Stage 3: Action** – A focus on targeted audiences that can influence and bring about change. | | | | |
| 1. To communicate that levels of negative test emotions can be significant for many children and young people.<br>2. Highlight research evidence on the consequences of test anxiety on young people.<br>3. To challenge the need for a testing culture in the light of the evidence.<br>4. To challenge the "*snowflake generation"* trope in the light of this research. | 1. School Leaders<br><br>**Research Findings A, B, E** | • There is a need to recognise that high test anxiety levels can affect a significant number of students – up to 89% of students depending on the point of time within the school year. Schools should recognise that for some there may be a facilitative effect, whereas for other this can be debilitating.<br>• Year 11 students' feelings around texting can increase levels of anxiety as well as feelings of shame and hopelessness. These appear to build-up over time as mock and final examinations get closer.<br>• Schools may wish to consider developing a whole school that takes an informing, recognising, understanding and emotional management approach.<br>• Where needed, schools may wish to offer individualised support for students who struggle with high levels of negative test emotions, including counselling, mentoring, or academic support to address the root causes of their emotional distress.<br>• Ensure a school culture that values learning and growth over test scores, helping to reduce the pressure and shame associated with exam performance. | • Local Authority briefings for schools.<br>• Local Authority headteacher association meetings.<br>• Articles for National College for School Leaders publications.<br>• Offer professional development workshops for schools' development days. | • Ongoing form September 2023 |
| | 1. Ofqual<br>2. Governmental Select Committees<br><br>**Research Findings C, D & E.** | • There is a need to recognise that high test anxiety levels can affect a significant number of students – up to 89% of students depending on the point of time within the school year.<br>• Research evidence indicates that high stress levels have negative health consequences for some young people. As such, we should look for alternatives to examinations or reduce their use as much as possible, e.g., modular models of assessment.<br>• That the constant evaluative culture in schools, encouraged by inspection and standards frameworks is a driver of stress for young people.<br>• That this is not a "*snowflake generation*" rather a "*test emotionally exhausted*" generation.<br>• There is a need to commission further research around the test anxiety and wider test emotions, specifically around the current construct validity of test anxiety.<br>• A current understanding of children and young people's perspectives and experiences of test anxiety is needed to inform test anxiety measures and good practice. | • Letters/article in the Times Educational Supplement.<br>• Online petitions.<br>• Examination board/Ofqual consultation responses.<br>• Relevant Governmental Select Committees. | • Ongoing form September 2023 |

**References**


Alberts, G, & Godlee, R. (2017, October 16). Sharing is caring: The ethical and moral obligation of research publication. New Scientist, 2854, 17-18.

Barkham, M., Hardy, G.E., & Mellor-Clark, J. (2010). Developing and delivering practice-based evidence: A guide for phycological therapies. Wiley-Blackwell,

Bazeley, P. (2021), Qualitative data analysis. London: Sage.

Cochrane, A.L. 1931-1971: a critical review, with particular reference to the medical profession. In: Teeling-Smith G, Wells N, editors. Medicines for the year 2000. London (UK): Office of Health Economics; 1979. 1-11.

Dunsmuir, S., Brown, E., Iyadurai, S., & Monsen, J. (2009). Evidence-based practice and evaluation: From insight to impact. Educational Psychology in Practice, 25(1), 53-70

Fallon, K., Woods, K., & Rooney, R. (2010). A discussion of the developing role of educational psychologists within Children's Services, Educational Psychology in Practice, 26:1, 1-23

Flick, U. (2014). An introduction to qualitative research. London: Sage

Frederickson, N. (2002). Evidence-based practice and educational psychology. London: Wiley

Godfrey, D., & Brown, C. [Eds.]. (2019). An ecosystem for research-engaged schools: Reforming education through research. London: Routledge.

Gough, D., Oliver, S., & Thomas, J. (2012). An introduction to systematic reviews. London: Sage.

Graham, I. D., Logan, J., Harrison, M. B., Straus, S. E., Tetroe, J., Caswell, W., & Robinson, N. (2006). Lost in knowledge translation: Time for a map? Journal of Continuing Education in the Health Professions, 26(1), 13-24.

Greenhalgh, T., & Wieringa, S. (2011). Is it time to drop the 'knowledge translation' metaphor? A critical literature review. Journal of the Royal Society of Medicine, 104(12), 501-509.

Harmsworth, S., & Turpin, S. (2000). Creating an Effective Dissemination Strategy. Bridging the Gap - Innovations Project, TQEF National Co-ordination Team.

Health and Care Professions Council. (2016). Practitioner psychologists. Health and Care

Professions Council Website. Available at http://doi.org/http://www.hpc-uk.org/assets/documents/10002963SOP_Practitioner_psychologists.pdf Accessed on 25th of April 2022.

Hoffman, A.J. (2021). The engaged scholar: Expanding the impact of academic research in today's world. Stanford, California: Stanford Press

Kelly, W. [Ed.]. (2022). The impactful academic: Building a research career that makes a difference. London: Emerald Publishing

Knight, C., & Lightowler, C. (2010). Reflecting on knowledge exchange: How effective is the Scottish approach? Evidence & Policy. Journal of Research, Debate and Practice, 6(4), 477-495.

Lane, D. A., Corrie, S. (2006). The modern scientist-practitioner: A guide to practice in phycology. London: Routledge

Lavis, J. N., Robertson, D., Woodside, J. M., McLeod, C. B., & Abelson, J. (2003). How can research organizations more effectively transfer research knowledge to decision-makers? The Milbank Quarterly, 81(2), 221-248.

McNiff, J. & Whitehead, J. (2011). All you need to know about action research, 2nd edition. London: Sage Publications.

Mellor, N., & Stoloff, M. (2017). Promoting the transfer of knowledge in educational psychology: An analysis of current practice and future directions. Journal of Educational and Psychological Consultation, 27(4), 457-476.

Nutley, S. M., Walter, I., & Davies, H. T. (2007). Using evidence: How research can inform public services. London: Policy Press.

Scott, A., Shaw, M., & Joughin, C. (2001). Finding the evidence: A gateway to the literature in child and adolescent mental health. London: Gaskell (Royal College of Psychiatrists).

Taylor, M. J. (2019). Knowledge transfer in educational psychology: An emerging agenda for practice and research. Educational Psychology in Practice, 35(3), 225-235

## Appendix 1 – Journal submission guideline for Applied Measurement in Education

**AUTHOR GUIDELINES**

**Active from: 01/03/2018**

(1) Submission

(2) Aims and Scope

(3) Manuscript Categories and Requirements

(4) Preparing the Submission

(5) Editorial Policies and Ethical Considerations

(6) Author Licensing

(7) Publication Process After Acceptance

(8) Post Publication

(9) Editorial Office Contact Details

1. SUBMISSION

Authors should kindly note that submission implies that the content has not been published or submitted for publication elsewhere except as a brief abstract in the proceedings of a scientific meeting or symposium.

Once the submission materials have been prepared in accordance with the Author Guidelines, manuscripts should be submitted online at https://mc.manuscriptcentral.com/pits

**Data protection**: By submitting a manuscript to or reviewing for this publication, your name, email address, and affiliation, and other contact details the publication might require, will be used for the regular operations of the publication, including, when necessary, sharing with the publisher (Wiley) and partners for production and publication. The publication and the publisher recognize the importance of protecting the personal information collected from users in the operation of these services, and have practices in place to ensure that steps are taken to maintain the security, integrity, and privacy of the personal data collected and processed. You can learn more at https://authorservices.wiley.com/statements/data-protection- policy.html.

**Preprint policy:** Wiley believes that journals publishing for communities with established pre-print servers should allow authors to submit manuscripts which have already been made available on a non-commercial preprint server. Allowing submission does not, of course, guarantee that an article will be sent out for review. It simply reflects our belief that journals should not rule out reviewing a paper simply because it has already been available on a non-commercial server. Please see below for the specific policy language.

However, Wiley also knows that the use of preprint servers is not universally accepted and that individual journals and/or societies may approach submission of preprints differently.
This journal will consider for review articles previously available as preprints on non-commercial servers such as ArXiv, bioRxiv, psyArXiv, SocArXiv, engrXiv, etc. Authors may also post the submitted version of a manuscript to non-commercial servers at any time. Authors are requested to

update any pre-publication versions with a link to the final published article. For help with submissions, please contact: support@scholarone.com

2. AIMS AND SCOPE

Psychology in the Schools is a peer-reviewed journal devoted to research, opinion, and practice. The journal, which is published ten times per year, welcomes theoretical and applied manuscripts, focusing on the issues confronting school psychologists, teachers, counsellors, administrators, and other personnel workers in schools and colleges, public and private organizations. Preference will be given to manuscripts that clearly describe implications for the practitioner in the schools.

3. MANUSCRIPT REQUIREMENTS

**Publication Charges:** Color figures. As outlined above, color figures may be published online free of charge, however the journal charges for publishing figures in color in print. Authors who supply color figures will be sent a Color Work Agreement once their accepted paper moves to the production process. If the color Work Agreement is not returned by the specified date figures will be converted to black and white for prin publication. Instructions on how to pay for the charges will be provided in the Color Work Agreement. For more detailed information on billing and payment, please click here https://authorservices.wiley.com/author-resources/Journal-Authors/open-access/article-publication-charges.html

**Practitioner Points:** Authors will need to provide no more than 3 'key points' that summarize the key messages of their paper to be published with their article. The key points should be written with a practitioner audience in mind. Reproduction of Copyright Material
If excerpts from copyrighted works owned by third parties are included, credit must be shown in the contribution. It is the author's responsibility to also obtain written permission for reproduction from the copyright owners. For more information visit Wiley's Copyright Terms & Conditions FAQ at http://exchanges.wiley.com/authors/faqs---copyright-terms--conditions_301.html

4. PREPARING THE SUBMISSION

Cover Letters: Cover letters are not mandatory; however, they may be supplied at the author's discretion.

Parts of the Manuscript: The manuscript should be submitted in separate files: title page; main text file; figures. Title Page

The title page should contain:

- A short informative title containing the major key words. The title should not contain abbreviations (see Wiley's best practice SEO tips);
- A short running title of less than 40 characters;
  The full names of the authors;
  The author's institutional affiliations where the work was conducted, with a footnote for the
- author's present address if different from where the work was conducted.
- Acknowledgments.
- Authorship
- Please refer to the journal's Authorship policy in the Editorial Policies and Ethical Considerations section for details on author listing eligibility.
  Acknowledgments
  Contributions from anyone who does not meet the criteria for authorship should be listed, with permission from the contributor, in an Acknowledgments section. Financial and material support should also be mentioned. Thanks to anonymous reviewers are not appropriate.
- Conflict of Interest Statement: Authors will be asked to provide a conflict of interest statement during the submission process. For details on what to include in this section, see the 'Conflict of Interest' section in the Editorial Policies and Ethical Considerations section below.

Submitting authors should ensure they liaise with all co-authors to confirm agreement with the final statement.

- Main Text File: As papers are double-blind peer reviewed, the main text file should not include any information that might identify the authors.The main text file should be presented in the following order:
    - ii. iii.
    - iv. v. vi.
- Title: The title should be short and informative. The main text of the manuscript should be divided into subsections for clarity. Authors are responsible for the accuracy of all written material.
- Abstract: A 250–350-word abstract communicating the essence of the paper is required. The abstract should succinctly and accurately describe the paper so that appropriate referees can be matched to the topic. Main Text/Copy: Title, abstract, and key words; Main text;
- References;
- Tables (each table complete with title and footnotes); Figure legends;
- Appendices (if relevant).

All copy should be typed double-spaced with one-inch margins. Microsoft Word 6.0 is preferred, although manuscripts prepared with any other microcomputer word processor are acceptable. LaTeX is usable but not preferred.

ALL MANUSCRIPTS MUST BE ACCOMPANIED BY AN RTF OR TXT FILE THAT CAN BE EDITED TO JOURNAL SPECIFICATIONS.

If you use software other than Microsoft Word, please provide a plain text file of the manuscript. Most applications support the export of plain text files. Please do not submit text files in Adobe PageMaker, Quark Xpress, or other desktop publishing software. If you prepared your manuscript with desktop publishing software, export the text to a word processing format. Refrain from complex formatting; the publisher will style your manuscript according to journal design specifications. Please make sure your word processing program's "fast save" feature is turned off. Please do not deliver files that contain hidden text. For example, do not use your word processor's automated features to create footnotes or reference lists.

The introduction should include a statement of the problem being addressed, why it is important, and to whom it is important. How is the study related to other work? Is it an extension? Major or minor? Is it a correction or difference of interpretation?

The conclusion should tell the reader clearly what the paper finds or demonstrates. It should be consistent with the objectives set forth in the introduction. It should describe the implications of the results for researchers, traders, policy makers, etc.

Notes: Notes and other short communications will be considered for publication as well as Comments on recent articles.

Appendix: Wherever possible, detailed mathematical analysis should be placed in an Appendix.
Footnotes: Expository footnotes should be cited in the text with a superscript Arabic number and typed double-spaced on a separate page at the end of the manuscript. When typeset, the footnotes will appear at the bottom of the page on which they are cited.
Figures:
Figures and supporting information should be supplied as separate files.
Keywords
psychology; schools; school psychologists; teachers; counselors; administrators

Main Text

• As papers are double-blind peer reviewed, the main text file should not include any information that might identify the authors.

• The journal uses British/US spelling; however, authors may submit using either option, as spelling of accepted papers is converted during the production process.
• Footnotes to the text are not allowed and any such material should be incorporated into the text as parenthetical matter.

References

References should be prepared according to the Publication Manual of the American Psychological Association (6th edition). This means in text citations should follow the author-date method whereby the author's last name and the year of publication for the source should appear in the text, for example, (Jones, 1998). The complete reference list should appear alphabetically by name at the end of the paper.

A sample of the most common entries in reference lists appears below. Please note that a DOI should be provided for all references where available. For more information about APA referencing style, please refer to the APA FAQ. Please note that for journal articles, issue numbers are not included unless each issue in the volume begins with page one.

Journal article:

Example of reference with 2 to 7 authors

Beers, S. R., & De Bellis, M. D. 2002. Neuropsychological function in children with maltreatment-related posttraumatic stress disorder The American Journal of Psychiatry159, 483–486. doi: 10.1176/appi.ajp.159.3.483

Ramus, F., Rosen, S., Dakin, S. C., Day, B. L., Castellote, J. M., White, S., & Frith, U. (2003). Theories of developmental dyslexia: Insights from a multiple case study of dyslexic adults. Brain, 126(4), 841– 865. doi: 10.1093/brain/awg076

Example of reference with more than 7 authors

Rutter, M., Caspi, A., Fergusson, D., Horwood, L. J., Goodman, R., Maughan, B., ... Carroll, J. (2004). Sex differences in developmental reading disability: New findings from 4 epidemiological studies. Journal of the American Medical Association, 291(16), 2007–2012. doi: 10.1001/jama.291.16.2007

Book edition:

Bradley-Johnson, S. (1994). Psychoeducational assessment of students who are visually impaired or blind: Infancy through high school(2nd ed.). Austin, TX: Pro-ed.

References should be typed double-spaced, in alphabetical order starting on a separate page following the manuscript. References should refer only to material listed within the text. Do not abbreviate journal names. Authors should review and verify references before manuscripts are submitted for consideration, because they alone are responsible for accuracy and completeness. Anthologies and collections must include names of editors and pages on which the reference appears. Books in a series must include series title and number/volume if applicable. Because of the large quantity of conference proceedings available, it is critical to give as much information as possible when citing references from proceedings. Please include the complete title of the meeting, symposium, etc. (do not abbreviate titles), and the city and dates of the meeting. If a proceeding has been published, please provide the editors' names, publisher, city, and year of publication, and pages on which the article appears.

Footnotes:

Footnotes should be placed as a list at the end of the paper only, not at the foot of each page. They should be numbered in the list and referred to in the text with consecutive, superscript Arabic numerals. Keep footnotes brief; they should contain only short comments tangential to the main argument of the paper and should not include references.

Tables:

Tables should be self-contained and complement, not duplicate, information contained in the text. They should be supplied as editable files, not pasted as images. Legends should be concise but comprehensive – the table, legend, and footnotes must be understandable without reference to the text. All abbreviations must be defined in footnotes. Footnote symbols: †, ‡, §, ¶, should be used (in that order) and *, **, *** should be reserved for P-values. Statistical measures such as SD or SEM should

be identified in the headings.

Figure Legends

Legends should be concise but comprehensive – the figure and its legend must be understandable without reference to the text. Include definitions of any symbols used and define/explain all abbreviations and units of measurement.

Figures

Although authors are encouraged to send the highest-quality figures possible, for peer-review purposes, a wide variety of formats, sizes, and resolutions are accepted.
Click here for the basic figure requirements for figures submitted with manuscripts for initial peer review, as well as the more detailed post-acceptance figure requirements.

Figures and artwork

Gray scale art & line art. As with tables (see above), figures should not be embedded in the text. Please cite the figure in the article text and provide a list of figure captions at the end of the manuscript, after the references. Figures must be numbered consecutively with arabic numerals. Figures are best submitted in tiff or eps (with preview) formats. Each figure should be in a separate file. Please do not submit proprietary graphics formats such as corel draw or adobe illustrator. Authors concerned with best quality printing should ensure that gray scale figures (e.g., screen shots, photos, or charts requiring shades of gray) are high resolution (above 300 dpi). Figures pasted directly from the web are low resolution (72 dpi). Bitmapped line art (made only of black & white lines—often simple charts or graphs) should be submitted at higher resolutions yielding 600-1200 dpi.
Authors are cautioned to provide lettering of graphs and figure labels that is large, clear, and open so that letters and numbers do not become illegible when reduced. Likewise, authors are cautioned that very thin lines and other fine details in figures may not successfully reproduce. Original figures should be created with these precautions in mind.

Data Citation

In recognition of the significance of data as an output of research effort, Wiley has endorsed the FORCE11 Data Citation Principles and is implementing a mandatory data citation policy. Wiley journals require data to be cited in the same way as article, book, and web citations and authors are required to include data citations as part of their reference list.

Data citation is appropriate for data held within institutional, subject focused, or more general data repositories. It is not intended to take the place of community standards such as in-line citation of GenBank accession codes.

When citing or making claims based on data, authors must refer to the data at the relevant place in the manuscript text and in addition provide a formal citation in the reference list. We recommend the format proposed by the Joint Declaration of Data Citation Principles:
[dataset] Authors; Year; Dataset title; Data repository or archive; Version (if any); Persistent identifier (e.g. DOI)
Additional Files

Appendices

Appendices will be published after the references. For submission they should be supplied as separate files but referred to in the text.

Supporting Information

Supporting information is information that is not essential to the article, but provides greater depth and background. It is hosted online and appears without editing or typesetting. It may include tables, figures, videos, datasets, etc.
Click here for Wiley's FAQs on supporting information.

Note: if data, scripts, or other artefacts used to generate the analyses presented in the paper are available via a publicly available data repository, authors should include a reference to the location of the material within their paper.

General Style Points

The following points provide general advice on formatting and style.

- Abbreviations: In general, terms should not be abbreviated unless they are used repeatedly, and the abbreviation is helpful to the reader. Initially, use the word in full, followed by the abbreviation in parentheses. Thereafter use the abbreviation only.
- Units of measurement: Measurements should be given in SI or SI-derived units. Visit the Bureau International des Poids et Mesures (BIPM) website for more information about SI units.
- Numbers: numbers under 10 are spelt out, except for: measurements with a unit (8mmol/l); age (6 weeks old), or lists with other numbers (11 dogs, 9 cats, 4 gerbils).
- Trade Names: Chemical substances should be referred to by the generic name only. Trade names should not be used. Drugs should be referred to by their generic names. If proprietary drugs have been used in the study, refer to these by their generic name, mentioning the proprietary name and the name and location of the manufacturer in parentheses.

Wiley Author Resources

Manuscript Preparation Tips: Wiley has a range of resources for authors preparing manuscripts for submission available here. In particular, we encourage authors to consult Wiley's best practice tips on Writing for Search Engine Optimization.

A video abstract can be a quick way to make the message of your research accessible to a much larger audience. Wiley and its partner Research Square offer a service of professionally produced video abstracts, available to authors of articles accepted in this journal. You can learn more about it by clicking here. If you have any questions, please direct them to videoabstracts@wiley.com. 5.

EDITORIAL POLICIES AND ETHICAL CONSIDERATIONS

Peer Review and Acceptance

The acceptance criteria for all papers are the quality and originality of the research and its significance to journal readership. Except where otherwise stated, manuscripts are double-blind peer reviewed. Paper Editing, Translation, and Formatting Support: Wiley Editing Services can greatly improve the chances of a manuscript being accepted. Offering expert help in English language editing, translation, manuscript formatting, and figure preparation, Wiley Editing Services ensures that the manuscript is ready for submission.

Video Abstracts will only be sent to review if the Editor-in-Chief determines that the paper meets the appropriate quality and relevance requirements. Wiley's policy on the confidentiality of the review process is available here.

Conflict of Interest

The journal requires that all authors disclose any potential sources of conflict of interest. Any interest or relationship, financial or otherwise that might be perceived as influencing an author's objectivity is considered a potential source of conflict of interest. These must be disclosed when directly relevant or

directly related to the work that the authors describe in their manuscript. Potential sources of conflict of interest include, but are not limited to: patent or stock ownership, membership of a company board of directors, membership of an advisory board or committee for a company, and consultancy for or receipt of speaker's fees from a company. The existence of a conflict of interest does not preclude publication. If the authors have no conflict of interest to declare, they must also state this at submission. It is the responsibility of the corresponding author to review this policy with all authors and collectively to disclose with the submission ALL pertinent commercial and other relationships.

Funding

Authors should list all funding sources in the Acknowledgments section. Authors are responsible for the accuracy of their funder designation. If in doubt, please check the Open Funder Registry for the correct nomenclature: https://www.crossref.org/services/funder-registry/
Funder arrangements (e.g. Wellcome Trust, RCUK, Austrian Science Fund and others)
There are separate arrangements with certain funding agencies governing reuse of this version as set forth at the following website: http://www.wiley.com/go/funderstatement.

Authorship

All listed authors should have contributed to the manuscript substantially and have agreed to the final submitted version.

Data Sharing and Data Accessibility

The journal recognizes the many benefits of archiving research data. The journal expects you to archive all the data from which your published results are derived in a public repository. The repository that you choose should offer you guaranteed preservation (see the registry of research data repositories at https://www.re3data.org/) and should help you make it findable, accessible, interoperable, and re- useable, according to FAIR Data Principles (https://www.force11.org/group/fairgroup/fairprinciples).

All accepted manuscripts are required to publish a data availability statement to confirm the presence or absence of shared data. If you have shared data, this statement will describe how the data can be accessed, and include a persistent identifier (e.g., a DOI for the data, or an accession number) from the repository where you shared the data. Authors will be required to confirm adherence to the policy. If you cannot share the data described in your manuscript, for example for legal or ethical reasons, or do not intend to share the data then you must provide the appropriate data availability statement. The journal notes that FAIR data sharing allows for access to shared data under restrictions (e.g., to protect confidential or proprietary information) but notes that the FAIR principles encourage you to share data in ways that are as open as possible (but that can be as closed as necessary).

Sample statements are available http://authorservices.wiley.com. If published, all statements will be placed in the heading of your manuscript.

Data Citation

Please also cite the data you have shared, like you would cite other sources that your article refers to, in your references section. You should follow the format for your data citations laid out in the Joint Declaration of Data Citation Principles, https://www.force11.org/datacitationprinciples:
[dataset] Authors; Year; Dataset title; Data repository or archive; Version (if any); Persistent identifier (e.g. DOI)

Publication Ethics

This journal is a member of the Committee on Publication Ethics (COPE). Note this journal uses iThenticate’s CrossCheck software to detect instances of overlapping and similar text in submitted manuscripts. Read Wiley’s Top 10 Publishing Ethics Tips for Authors here. Wiley’s Publication

Ethics Guidelines can be found here.

ORCID
As part of the journal's commitment to supporting authors at every step of the publishing process, the journal requires the submitting author (only) to provide an ORCID ID when submitting a manuscript. This takes around 2 minutes to complete. Find more information here.

6. AUTHOR LICENSING

If a paper is accepted for publication, the author identified as the formal corresponding author will receive an email prompting them to log in to Author Services, where via the Wiley Author Licensing Service (WALS) they will be required to complete a copyright license agreement on behalf of all authors of the paper.

Authors may choose to publish under the terms of the journal's standard copyright agreement, or OnlineOpen under the terms of a Creative Commons License.

General information regarding licensing and copyright is available here. To review the Creative Commons License options offered under OnlineOpen, please click here. (Note that certain funders mandate a particular type of CC license be used; to check this please click here.)
Self-Archiving Definitions and Policies: Note that the journal's standard copyright agreement allows for self-archiving of different versions of the article under specific conditions. Please click here for more detailed information about self-archiving definitions and policies.
Open Access fees: Authors who choose to publish using OnlineOpen will be charged a fee. A list of Article Publication Charges for Wiley journals is available here.
Funder Open Access: Please click here for more information on Wiley's compliance with specific Funder Open Access Policies.

7. PUBLICATION PROCESS AFTER ACCEPTANCE

Accepted Article Received in Production

When an accepted article is received by Wiley's production team, the Wiley Author Services
Accepted Articles

The journal offers Wiley's Accepted Articles service for all manuscripts. This service ensures that accepted 'in press' manuscripts are published online shortly after acceptance, prior to copy-editing or typesetting. Accepted Articles are published online a few days after final acceptance, appear in PDF format only, and are given a Digital Object Identifier (DOI), which allows them to be cited and tracked. After publication of the final version article (the article of record), the DOI remains valid and can still be used to cite and access the article.

Proofs

Once the paper is typeset, the author will receive an email notification with full instructions on how to provide proof corrections.

Please note that the author is responsible for all statements made in their work, including changes made during the editorial process – authors should check proofs carefully. Note that proofs should be returned within 48 hours from receipt of first proof.

Accepted Articles

This journal offers rapid publication via Wiley's Accepted Articles Service. Accepted Articles are published on Wiley Online Library within 5 days of receipt, without waiting for a copyright agreement

(they are protected under a general copyright statement).

Early View

The journal offers rapid publication via Wiley's Early View service. Early View (Online Version of Record) articles are published on Wiley Online Library before inclusion in an issue. Note there may be a delay after corrections are received before the article appears online, as Editors also need to review proofs. Before we can publish an article, we require a signed license (authors should login or register with Wiley Author Services. Once the article is published on Early View, no further changes to the article are possible. The Early View article is fully citable and carries an online publication date and DOI for citations.

8. POST PUBLICATION

Access and Sharing
When the article is published online:
The author receives an email alert (if requested).
The link to the published article can be shared through social media.
The author will have free access to the paper (after accepting the Terms & Conditions of use, they can view the article). corresponding author will receive an email asking them to login or register with
The author will be asked to sign a publication license at this point.

For non-open access articles, the corresponding author and co-authors can nominate up to ten colleagues to receive a publication alert and free online access to the article.

Promoting the Article

To find out how to best promote an article, click here.

Measuring the Impact of an Article

Wiley also helps authors measure the impact of their research through specialist partnerships with Kudos and Altmetric.

Archiving Services

Portico and CLOCKSS are digital archiving/preservation services we use to ensure that Wiley content will be accessible to customers in the event of a catastrophic event such as Wiley going out of business or the platform not being accessible for a significant period of time. Member libraries participating in these services will be able to access content after such an event. Wiley has licenses with both Portico and CLOCKSS, and all journal content gets delivered to both services as it is published on Wiley Online Library. Depending on their integration mechanisms, and volume loads, there is always a delay between content being delivered and showing as "preserved" in these products.

9. EDITORIAL OFFICE CONTACT DETAILS
David E. McIntosh, Ball State University jnlpsychscho@bsu.edu

Author Guidelines updated 5/22/19

**Appendix 2: Letter to journal editor.**

Anthony Collins, Trainee Educational Psychologist
Doctorate in Educational & Child Psychology
School of Environment, Education and Development
Ellen Wilkinson Building
University of Manchester
Oxford Road
Manchester M13 9PL

Dear Mr Geisinger,

I am writing to submit a paper titled "Testing Times: A Longitudinal Study of Test-Related Emotions" for consideration in Applied Measurement in Education. I am aware that the journal typically adheres to strict length requirements for submissions; however, I kindly request your consideration to make an exception in this case as we believe that the depth and breadth of our research necessitate a longer article to effectively convey the comprehensive results and valuable insights we have obtained.

Our study spanned ten months and followed a large cohort of participants, allowing us to examine a range of test emotion variables over an extended period. The longitudinal design enabled us to capture nuanced changes, uncover critical patterns, and address complex research questions. By closely tracking our participants, we have been able to generate substantial data that offers significant contributions to the field.

The length of our paper, which currently exceeds the prescribed limit of 35 pages, is a consequence of the comprehensive nature of our study. To condense our findings into a shorter article would inevitably result in a loss of essential details, limiting the reader's understanding of the methodology, results, and implications of our research. By granting an exception to the length requirement, we would ensure that our paper can do justice to the depth and significance of our findings.

We understand that accommodating a longer paper may require additional editorial effort, and we are willing to work closely with you to ensure that the manuscript conforms to the journal's formatting guidelines while retaining its essential content. We are also open to suggestions regarding sections that could be condensed or revised to enhance readability without compromising the core message and findings.

We firmly believe that our research holds significant value to the academic community and that its inclusion in Applied Measurement in Education would greatly contribute to the advancement of knowledge in the field. We appreciate your time and consideration of our request, and we are confident that our work aligns with the objectives and scope of your esteemed journal.

Yours sincerely,

Anthony Collins

## Appendix 3: Test Anxiety Analytic Framework

<table>
<tr><td colspan="5">Name of Measure:</td></tr>
<tr><td colspan="2">Author:</td><td>Publisher:</td><td>Age Range:</td><td>Date of Norms:</td></tr>
<tr><td colspan="5">Standardisation Sample:</td></tr>
<tr><td colspan="5">Purpose/Function:</td></tr>
<tr><td colspan="5">Description:</td></tr>
<tr><td>Admin Time:</td><td>Reliability:</td><td>Validity:</td><td>How scales are reported:</td><td>Confidence interval:</td></tr>
<tr><td colspan="5">Subtest Information/Psychological construction: (What factors are included, what is the rationale for inclusion)</td></tr>
<tr><td colspan="5">Inclusion of the views of young people?</td></tr>
<tr><td colspan="5">Comment:</td></tr>
</table>

## Appendix 4: Participant Information Sheet


The University of Manchester

### Participant Information Sheet (PIS)

You are being invited to take part in a research study that forms one of the assessment requirements for the Doctorate in Educational and Child Psychology at the University of Manchester. The aim of this piece of research is to explore the young peoples' emotions around tests and examinations across Year 11. The findings from this research project may be used to inform the researcher's doctoral thesis.

Before you decide whether to take part, it is important for you to understand why the research is being conducted and what it will involve. Please take time to read the following information carefully before deciding whether to take part and discuss it with others if you wish. Please ask if there is anything that is not clear or if you would like more information. Thank you for taking the time to read this.

**Who will conduct this research?**

Anthony Collins (Trainee Educational Psychologist)
School of Environment, Education and Development (SEED),
Ellen Wilkinson Building,
Oxford Road,
Manchester,
M13 9PL.

**Title of the Research**

*'Testing Times: A Longitudinal Study of Year 11 GCSE students Test-Related Emotions'*

**What is the purpose of the research?**

This research aims to explore students test-related emotions across the academic year (September – August) as they study towards, take, and await the results of their GCSE examinations. The aim is to use the data to improve the support offered to students by identifying points in time when additional pastoral support or targeted interventions maybe needed.

**Why have I been chosen?**

You have been chosen as a student who is in Year 11 and working towards GCSE examinations in the May/June 2021 exam period.

**Who has reviewed the research project?**

The research project has followed Ethical Practice Policy and Guidance set by the Manchester Institute of Education and has been reviewed by supervisor, Prof. Kevin Woods.

**Will the outcomes be published?**

This exploratory study provides information towards a subsequent doctoral thesis and forms an assignment which will be submitted towards the researcher's doctoral training. It is possible therefore that this research may be published on completion of the doctoral thesis in 2023. This research may also be shared with the commissioning Educational Psychology Service since it is commissioned by them.

**What would I be asked to do if I took part?**

Every eight weeks across the academic year, you will be asked to complete an online questionnaire that asks what kind of tests or examinations you have had in the past eight-weeks and how you have felt before, during and after those tests.

**Will I be compensated for taking part?**

Participation is voluntary and no financial compensation is offered for participation.

**What happens if I do not want to take part or change my mind?**

It is up to you to decide whether you wish to take part. You will be asked to click a check box to indicate you consent to take part. If you decide to take part, you can withdraw from the research at any time prior to completing the questions. Following participation, you can have your data withdrawn from the study up to ten working days following participation.

**Data Protection & Confidentiality**

Please also note that individuals from The University of Manchester or regulatory authorities may need to look at the data collected for this study to make sure the project is being carried out as planned. We will not collect any identifiable data that might identify you. All individuals involved in auditing and monitoring the study will have a strict duty of confidentiality as a research participant.

**Contact Details**

**What if you want more information?**

Contacts for further information:
Researcher: Anthony Collins
Email: anthony.collins-2@postgrad.manchester.ac.uk

Supervisor: Prof. Kevin Woods
Email: kevin.a.woods@manchester.ac.uk

**What if I have a complaint?**

You are welcome to contact the researcher or supervisor in the first instance. If there are any issues regarding this research that you would prefer not to discuss with members of the research team, please contact the Research Practice and Governance Coordinator by either writing to: The Research Practice and Governance Co-ordinator, Research Office, Christie Building, The University of Manchester, Oxford Road, Manchester, M13 9PL or by emailing: Research-Governance@manchester.ac.uk

If you wish to make a formal complaint to someone independent of the research team or if you are not satisfied with the response you have gained from the researchers in the first instance then please contact: The Research Ethics Manager, Research Office, Christie Building, The University of Manchester, Oxford Road, Manchester, M13 9PL, by emailing: research.complaints@manchester.ac.uk or by telephoning 0161 275 2674.

If you wish to contact us about your data protection rights, please email dataprotection@manchester.ac.uk or write to The Information Governance Office, Christie Building, The University of Manchester, Oxford Road, M13 9PL at the University and we will guide you through the process of exercising your rights.
You also have a right to complain to the Information Commissioner's Office about complaints relating to your personal identifiable information Tel: 0303 123 1113

**Participant Consent**

By checking the consent box you are indicating the following:

1. I confirm that I have read the above information sheet (Version 1, 01/06/2021) for the above study and have had the opportunity to consider the information and ask questions and had these answered satisfactorily.
2. I understand that my participation in the study is voluntary and that I am free to withdraw at any time prior to submitting your responses without giving a reason and without detriment to myself. I understand that it will not be possible to remove my data from the project once it has been submitted and forms part of the data set.
3. I understand that the data will not be confidential as no personal or identifying information with gathered.
4. I agree to any of the anonymous data being passed to other researchers.
5. I agree that any data collected may be published in anonymous form in academic books, reports or journals.
6. I agree to take part in this study.

## Appendix 5: Qualtrics Survey


The following questions are about the feelings you may have BEFORE taking a test or exam. Please indicate how you have felt, typically, before take a test or or exam over the past two months.

(1=Not like me at all, 2=A little like me, 3=Sometimes like me, 4=Often like me, 5= Very much like me.)

1 2 3 4 5

1. I look forward to tests and exams.

2. I am optimistic that everything will work out fine.

3. I am so proud of my preparation that I want to start the exam now.

4. Before the exam I feel nervous and uneasy.

5. I look forward to demonstrating my knowledge.

6. Because I enjoy preparing for the test, I'm motivated to do more than I need to.

7. Because I look forward to being successful, I study hard.

8. Before taking the exam, I feel excited.

9. I have great hope that my abilities will be good enough.

10. I'm quite confident that I have prepared enough for the exam.

11. I think about my exam optimistically.

12. I start studying for the exam with great hope and expectation.

13. My confidence motivates me to prepare well.

14. I get angry over time pressures which don't leave enough time to prepare.

15. I get angry about the amount of material I need to know.

16. I worry whether I have studied enough.

17. I worry whether the test will be too hard.

18. I get so nervous I wish I could not do the exam.

19. I feel sick to my stomach.

20. I can't even think about how embarrassing it would be to fail the exam.

21. I get depressed because I feel I don't have much hope for the exam.

22. I have lost all hope that I have the ability to do well on the exam.

23. I feel so defeated that I can't start doing anything.

24. I'd rather not write the test because I have lost all hope.

25. Feeling hopeless means I have no energy.

→

MANCHESTER 1824
The University of Manchester

The following questions are about the feelings you may have DURING taking a test or exam. Please indicate how you have felt, typically, during a test or or exam over the past two months.

(1=Not like me at all, 2=A little like me, 3=Sometimes like me, 4=Often like me, 5= Very much like me.)

1 2 3 4 5

1. I enjoy taking tests and exams.

2. I am happy that I can cope with tests and exams.

3. For me the test is a challenge that is enjoyable.

4. I am very confident about tests and exams.

5. Hoping to succeed, I'm motivated to put in a lot of effort.

6. I think that I can be proud of my knowledge.

8. Because I'm proud of what I know I want to do my best in the test.

8. During tests I get angry.

9. I think the questions are unfair.

10. I am very nervous about tests and exams.

11. I feel panicky when writing exams.

12. I worry whether I will pass the exam.

13. I get nervous. I can't wait for the exam to be over.

14. I'm so anxious that I'd rather be anywhere else.

15. At the beginning of the test, my heart starts pounding.

16. My hands get shaky.

17. During an exam I feel humiliated.

18. I am ashamed that I have not prepared well enough.

19. I get embarrassed because I can't answer the questions correctly.

20. I get embarrassed and I want to leave.

21. Because I feel ashamed my pulse races.

22. I feel hopeless.

23. I have given up believing that I can answer the questions correctly.

24. I start to think that no matter how hard I try I won't succeed on the test or exam.

25. I start to realise that the questions are too difficult for me.

26. I feel like giving up.

27. I feel so defeated that I have no energy.

←

→

MANCHESTER 1824
The University of Manchester

The following questions are about the feelings you may have AFTER taking a test or exam. Please indicate how you have felt, typically, after a test or or exam over the past two months.

(1=Not like me at all, 2=A little like me, 3=Sometimes like me, 4=Often like me, 5= Very much like me.)

1 2 3 4 5

1. My heart beats faster with joy.

2. I am glowing with joy.

3. I am very satisfied with myself.

4. I am proud of myself.

5. To think about my success makes me feel proud.

6. I am proud of how I did in the exam.

7. When I get the test results back, my heart beats with pride.

8. After the exam I feel ten feet taller because I'm so proud.

9. I walk out of the exam with the look of a winner on my face.

10. I feel relief.

11. I fell freed.

12. I feel very relieved.

13. The worry in my stomach has gone.

14. I finally can breath easy again.

15. I can finally laugh again.

16. I am fairly annoyed.

17. I get angry about the teachers' grading standards.

18. I wish I could tell the teacher off.

19. I wish I could freely express my anger.

20. My anger makes the blood rush to my head.

21. I get so angry, I start feeling hot and flushed.

22. I feel ashamed.

23. My marks embarrass me.

24. When I get a bad mark I would prefer not to face my teacher again.

25. When others find out about my poor marks I start to blush.

←

→

MANCHESTER 1824
The University of Manchester

If you feel worried or upset about any of the issues raised in these questions a good first step would be to talk with a trusted adult such as a teacher, parents or carers. They are best placed to help you quickly find solutions to your concerns.

Childline offer advice and support about feeling and examination stress. You can find information here: https://www.childline.org.uk/info-advice/your-feelings/anxiety-stress-panic/coping-with-stress/

If you have a worry and you feel you need to talk with someone you can contact Childline directly on: 0800 1111

←

→