

### **Abstract**

Standardized testing is an integral part of the English and American education systems. However, the use of high-stakes testing has unintended consequences, one of which is test anxiety. Over the last 50 years, increased attention has been directed to developing tools to identify students experiencing test anxiety. However, many test anxiety instruments have been created for research purposes rather than use within school decision-making and lack evidence for interpretation. The purpose of the current study was to support the use of the Multidimensional Test Anxiety Scale (MTAS) in applied settings by using a Latent Profile Analysis to identify respondent groups to support score interpretation. Participants included 918 secondary students in England.

**Keywords:** test anxiety, latent profile analysis, decision-making

## Interpretation Evidence for the Multidimensional Test Anxiety Scale: A Brief Report

### **Introduction**

Academic anxiety is a comprehensive term for the various types of anxieties that students may experience in the school environment (Cassady, 2010). Test anxiety is one such academic anxiety and is defined as the changes in emotion and physiology resulting from an individual's perception of the consequences of a test or exam (Zeidner, 1998). Test anxiety occurs when an individual appraises an evaluative situation (i.e., test) as threatening and is often accompanied by worry and heightened physiological reactions. Students who experience test anxiety may perform poorly in exams due in part to the disruptive nature of test anxiety, such as devoting cognitive resources to non-test-related tasks (Angelidis A. et al., 2019). Test anxiety has been associated with poor academic performance since research into the subject began in the 1950s (Sarason & Mandler, 1952). A meta-analysis conducted by von der Embse and colleagues (2018) found a negative relationship between test anxiety and exam performance, grade point average (GPA) and standardized tests. Increases in test anxiety were associated with decreased scores across all three (von der Embse et al., 2018).

In America and the United Kingdom standardized tests, like the National Curriculum Tests and the Florida State Assessments, are used to evaluate students, teachers, schools and school districts, thus students are assessed at several points throughout their education. The scores on the standardized tests are used to determine whether students attain a high school diploma and can access post-secondary education (Segool et al. 2014), and they are often used to support or make educational decisions, like retention. A student may experience greater test anxiety from these standardized assessments than typical classroom exams because of the higher stakes associated with them (Segool et al., 2013).

Given the prominent role of standardized assessments in evaluating student academic performance and facilitating further educational attainment (i.e., university admission), it becomes essential for schools to identify students who may need additional support to reduce the likelihood of high test anxiety. Several test anxiety assessment tools have been developed over the last two decades, including the most widely used assessment, the Test Anxiety Index (TAI: Spielberger, 1980). However, there are several limitations to these tools. First, many test anxiety instruments were created for research purposes without the necessary evidence for use in applied decision-making situations. More specifically, evidence for *interpretation* (Kane, 2013) is an important consideration in developing an assessment such that assessment scores can meaningfully differentiate between groups. A second limitation is that several frequently used test anxiety scales are not reflective of modern advancements in theoretical conceptualizations of test anxiety.

The Multidimensional Test Anxiety Scale (MTAS; Putwain et al., 2020) was created to address the aforementioned limitations of several of the most frequently used test anxiety scales. The MTAS components included the cognitive aspects of worrying, thoughts around failure, and cognitive interference, and the autonomic aspects of physiological indicators and tension (Lowe et al., 2008; Zeidner & Matthews, 2005). To date, MTAS research has examined the factor structure and relation to distal academic performance and similar student report assessments (Putwain et al., 2020). Additionally, MTAS research has used a new sample to confirm the factor structure, evaluate measurement invariance, determine the internal consistency of the factors, and identify cut scores (von der Embse et al., 2021). The present study sought to add evidence for the applied use of the MTAS, by determining respondent groups via a Latent Profile Analysis.

There are various methods used to determine cut scores or standards for assessments. The two main types that these methods fall under are person centred analyses or variable centred analyses. In a previous study, cut scores of 58 and 60 in MTAS Total score were created for the MTAS compared to a panic and an anxiety scale using ROC (von der Embse, et al., 2021). The main advantage of ROC is the fact that it reports the changes in specificity and sensitivity across different cut of scores. However, that freedom to determine thresholds and cut scores is also a weakness. The level of specificity and sensitivity is dependent on the researcher and is thus subjective. Additionally, as a variable centred analysis it assumes that the sample is homogenous and focuses on the variables.

This study examined classification thresholds through a Latent Profile Analysis which is a person centred analysis based on the principle that responses to items in the assessment form distinct and mutually exclusive subgroups called latent profiles. This method can be used to determine both the number of classification groups formed through the assessment as well as the standards for inclusion in each of these classifications. Thus, it will expand on the process already started in von der Embse et al. (2021) to increase the usability of the MTAS by providing classification standards.

### **Methods**

There were 918 students included from an existing dataset, including 217 who self-identified as male and 694 who self-identified as female, while seven declined to disclose their gender. All participants were secondary school students from eight schools in the United Kingdom, two of which were girls' schools, which may account for the relatively higher number of females in the sample. The mean age of this sample was 15.76 years old. The grades in this sample were between Year 10 and Year 13. The racial/ethnic breakdown of this sample was 3%

Asian, 5% Black, 87% White, and 2% multiracial. Additionally, 15% of the sample were eligible for free school meals. There were no missing data in the dataset.

### **Measures**

**MTAS.** The Multidimensional Test Anxiety Scale (MTAS: Putwain & von der Embse, 2019) was developed to measure the multiple components of test anxiety. Factors include cognitive interference, worry, physiological indicators, and tension. The MTAS consists of 16 items and uses a Likert scale from 1 – “Strongly Disagree” to 5 “Strongly Agree” with 4 items for each of the four constructs listed above. The MTAS has supportive evidence for internal consistency, factorial validity, predictive validity, measurement invariance, and the creation of cut scores (Putwain et al., 2020; von der Embse et al., 2020). The MTAS has positive relationships with student mental health ( $r_s = .13$  to  $.46$ ) and negative relationships with academic performance ( $r_s = .01$  to  $.41$ ) and well-being ( $r_s = .01$  to  $.41$ ; Putwain et al., 2020).

**GAD.** Generalized Anxiety Disorder (GAD) was measured using the six-item Generalized Anxiety subscale from the Revised Children’s Anxiety and Depression Scale (RCADS: Chorpita et al., 2005). The subscale consisted of a 4-point scale from 0 (Never) to 3 (Always) whereby a higher score indicated greater anxiety. Subscales from the RCAD have good psychometric properties including convergent validity ( $r_s = .52$  to  $.74$  for Generalized Anxiety) and internal consistency ( $\alpha_s = .84$  to  $.88$  for GAD) (Donnelly et al., 2019).

### **Data Collection Procedure**

Data were collected from eight secondary schools. Once IRB approval was obtained, the eight schools were enrolled in the study, and consent was solicited from the principal, students, and parents of students under 18 and considered minors. The MTAS was then administered during a ‘free’ period in the students’ timetables when instruction was not being provided. The

teachers who administered the MTAS followed a script whereby they informed the students that they were not being tested, but to fill out the questionnaire honestly.

### **Data Analysis**

**Reliability.** The internal consistency of the four factors of the scale was examined. Cronbach's alpha and McDonald's omega were calculated for the different factors as well as the overall scale, with values between .7 and .9, indicating acceptable to exceptional internal consistency (Tavakol & Dennick, 2011).

**Latent Profile Analysis.** LPA was used to identify patterns of risk among the four factors, cognitive interference (CI), worry (W), physiological indicators (PI) and tension (T), and overall test anxiety (OTA). Based on the literature, the authors predicted that the classifications would be low, moderate, and high (von der Embse et al., 2014). However, to ensure that the ideal number of profiles was selected for the model, a two-profile model was run first, followed by a three-profile, a four-profile, and a five-profile model. The analysis stopped at a five-profile model, given the potential challenges in the interpretation and usability of six or greater profiles. Once a model was chosen, we used a multinomial logistic regression procedure to examine the association between the profile membership and measures of General Anxiety Disorder (GAD), using the R3STEP in Mplus 8.1 (Muthén & Muthén, 2017). Examination of the association between the profile membership to GAD was completed to provide validity evidence to the cluster solutions (3-step approach).

### **Results**

Table 1 displays the descriptive statistics for the items and scales of the MTAS. There were no variables with skewness greater than three or any with kurtosis greater than 10, which indicates that the data were approximately normally distributed (Chou & Bentler, 1995). The

correlations between the subscales ranged from 0.491 – 0.763, which indicates that the relationship between the subscales was moderate to large. The strongest correlations were between worry and tension ( $r=0.763$ ) and tension and physiological indicators ( $r=0.732$ ).

The AIC, BIC, Sample-adjusted BIC, LMR-LRT, adjusted LMR-LRT, Bootstrap LRT, and entropy were all used for model evaluation (see Table 2). The five-profile model was the best fit for the data (AIC = 17784.367, BIC = 17919.389, Sample-adjusted BIC = 17830.464); however, profile one only accounted for three percent of the sample and was not meaningful or practically useful for classification purposes. The four profile model had the next best model fit. Entropy was then considered whereby a criterion score closer to one indicated that profiles were distinct and the assignment of individuals to these profiles were accurate.

The interpretation of the LPA data included fit statistics along with the guiding test anxiety theory and the ultimate goal of identifying profiles that will aid in the interpretation of MTAS scores. The four profile model exhibited better fit (AIC = 17925.013, BIC = 18035.924, Sample-adjusted BIC = 17962.879, LMR-LRT = 186.953,  $p=.0429$ , Adjusted LMR-LRT = 181.628,  $p=.0458$ , Bootstrap LRT = 186.953,  $p=.000$ ). While the two profile model had the best entropy, it as well as the three profile and five profile models, were not the overall best fit based on the totality of the statistics. Thus the four profile model was selected.

Though previous research identified three profiles, four profiles may still allow for interpretation in school settings. Schools typically have limited resources to support students and the four profile model allows school personnel to quickly identify students experiencing the highest levels of test anxiety and provide the resources needed without overwhelming their systems. The four profile model also allows schools with more resources to identify the students who are not experiencing the highest levels of test anxiety but may still need support with

managing their lower levels of anxiety. Table 3 shows the mean scores and sample percentages of the profiles in each model. Profile Two was labeled low anxiety (average total score of 34). Profile One was labeled average anxiety (average score of 49). Profile Four was labeled above average anxiety (average score of 59). Profile Three was labeled high test anxiety because the average scores across the subscales were very high, ranging from 16-19 out of 20 and the average total score was also very high at 71 out of 80.

The multinomial logistic regression was run with profile 3 (high test anxiety) as the reference group. The unstandardized regression coefficients for the GAD indicated that the students with lower GAD symptoms were more likely to belong to the low test anxiety group ( $b = -0.578$ ,  $SE = .09$ ,  $p < .001$ ; odds ratio = xxx) and the average test anxiety group ( $b = -0.251$ ,  $SE = .07$ ,  $p < .001$ ; odds ratio = xxx) than the high test anxiety group.

### **Discussion**

Test anxiety has been linked with a number of negative outcomes, including low student performance in high-stakes tests (von der Embse et al., 2018), with significant downstream consequences such as grade retention or denial of university admission. There are a number of standardized tests that students take throughout their education, like the National Curriculum Tests, and these test scores are used to evaluate students' academic performance as well as the effectiveness of teachers and schools (Segool et al. 2014). Test anxiety impacts how well students do on these exams, so identifying and addressing test anxiety should be a high priority in schools. However, test anxiety must first be reliably identified to facilitate early intervention support. Test anxiety assessment tools are one of the most systematic ways to identify students experiencing test anxiety. The MTAS was developed to address limitations of existing tools, however additional research was needed to support use in school settings.

Interpretation and use include all the steps between administering an assessment and using its results to make decisions (Kane, 2013). The primary aim of the present study was to increase the usability of the MTAS in schools and other practical settings and provide guidelines for the interpretation of scores. Previous MTAS research identified cut scores using a variable-centered analytical approach with an external criterion (Putwain et al., 2020). This study took a different approach via LPA, which is a person-centered approach. An LPA offers a unique benefit by identifying clusters of responses with common attributes and based on theory (Laursen & Hoff, 2006).

Results indicated that a four-profile model was consistent with the fit statistics and test anxiety theory. Upon examination of the four profile model, the following descriptions were used including low test anxiety (average total score=34/80), average test anxiety (average total score=49/80), above average test anxiety (average total score=59/80), and high test anxiety (average total score=71/80). In the 3-step analysis, higher GAD symptoms were associated with a higher likelihood of belonging to the high test anxiety group, compared to the low test anxiety and average test anxiety groups. These four categories may be useful for schools and other stakeholders using the MTAS. The four profile model may allow school personnel to prioritize service for those students experiencing the high and above average levels of test anxiety. Approximately 20% of the sample endorsed high levels of anxiety which may indicate a substantial number of students in need of intervention support. These students experiencing high levels of test anxiety would likely benefit from individualized support. Additionally, many students (43%) endorsed above average levels of test anxiety and could potentially benefit from brief, small group interventions.

While there are strengths to the current study, there are some limitations. First, the present study utilized an existing dataset. This prevented modifications to how these data were collected, the diversity of the sample, and what types of data were collected. A second limitation included the representativeness of the sample, which may limit the generalization of results. Additional research is needed with a more diverse population of students from across different cultures and countries to examine the measurement invariance of the MTAS. Such a study would increase the population of students for whom the MTAS can be employed. Further research could also examine the number of latent profiles that emerge with different samples of students. It will be important to evaluate the membership within these profiles as potentially and differentially predictive of important distal outcomes such as academic achievement. Additionally, test anxiety is often comorbid with learning disorders and ADHD. Further research is needed to understand the relationship between test anxiety and these disorders and how they may interact with each other. Students with learning disorders and ADHD may already receive supports in the school setting but if they are also experiencing test anxiety additional supports may be needed to address this concern. Furthermore, future research should address treatment types and intensity to support the direction of limited resources. The four profile classification presented in this study separated scores into low, average, above average and high test anxiety. More research needs to be completed to investigate treatment for test anxiety across these different profiles. Lastly, research could be conducted to evaluate the stability of the profiles and how students may change over time through latent transition analysis. These further analyses would add support to the identified profiles and give additional information to users of the MTAS on the needs of students identified with higher levels of test anxiety. In an educational landscape that is increasingly reliant upon standardized testing for decision-making purposes,

understanding test anxiety and its impacts on test performance is crucial. The further development and validation of test anxiety assessment tools will be important to increase usability and, ultimately, decision-making.

### References

- Angelidis, A., Solis, E., Lautenbach, F., van der Does, W., & Putman, P. (2019). I'm going to fail! Acute cognitive performance anxiety increases threat-interference and impairs WM performance. *PloS One*, *14*(2), e0210824.
- Cassady, J.C. (2010). *Anxiety in Schools: The Causes, Consequences, and Solutions for Academic Anxieties* (2nd ed.) Peter Lang.
- Chorpita, B. F., Moffitt, C. E., & Gray, J. (2005). Psychometric properties of the Revised Child Anxiety and Depression Scale in a clinical sample. *Behaviour research and therapy*, *43*(3), 309-322.
- Chou, C. P., & Bentler, P. M. (1995). Estimates and tests in structural equation modeling. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 37–55). Sage Publications, Inc.
- Kane, M. T. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement*, *50*(1), 1-73. <https://doi.org/10.1111/jedm.12000>
- Laursen, B., & Hoff, E. (2006). Person-centered and variable-centered approaches to longitudinal data. *Merrill-Palmer Quarterly* *52*(3), 377-389.
- Lowe, P. A., Grumbein, M. J., & Raad, J. M. (2011). Examination of the Psychometric Properties of the Test Anxiety Scale for Elementary Students (TAS-E) Scores. *Journal of Psychoeducational Assessment*, *29*(6), 503–514.
- Lowe, P.A., Lee, S.W., Witteborg, K.M., Prichard, K.W., Luhr, M.E., Cullinan, C.M., Janik, M., (2008). The Test Anxiety Inventory for Children and Adolescents (TAICA): examination of the psychometric properties of a new multidimensional measure of test anxiety among

- elementary and secondary school students. *Journal of Psychoeducational Assessment* 26(3), 215–230. <http://dx.doi.org/10.1177/0734282907303760>.
- Muthén, L.K., & Muthén, B.O. (2017). *Mplus 8.1* [Computer software]. Los Angeles, CA, Muthén & Muthén
- Putwain, D. W., von der Embse, N., Rainbird, E. C., & West, G. (2020). The development and validation of a new Multidimensional Test Anxiety Scale (MTAS). *European Journal of Psychological Assessment* . <https://doi.org/10.1027/1015-5759/a000604>
- Sarason, S. B., & Mandler, G. (1952). Some correlates of test anxiety. *The Journal of Abnormal and Social Psychology*, 47(4), 810–817. <https://doi.org/10.1037/h0060009>.
- Segool, N., Carlson, J., Goforth, A., von der Embse, N., & Barterian, J. (2013). Heightened test anxiety among young children: Elementary school students' anxious responses to high-stakes testing. *Psychology in the Schools*, 50, 489-499.
- Segool, N., von der Embse, N.P., Mata, A., Gallant, J., (2014). Cognitive-behavioral model of test anxiety in a high-stakes context: An exploratory study. *School Mental Health* 6(1), 50–61. <http://dx.doi.org/10.1007/s12310-013-9111-7>.
- Spielberger, C. D. (1980). The test anxiety inventory. *Psychology*, 4(6A).
- Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International journal of medical education*, 2, 53. doi: 10.5116/ijme.4dfb.8dfd
- von der Embse, N., Jester, D., Roy, D., & Post, J. (2018). Test anxiety effects, predictors, and correlates: A 30-year meta-analytic review. *Journal of Affective Disorders*, 227, 483-493. <https://doi.org/10.1016/j.jad.2017.11.048>
- von der Embse, N., Kim, E., Jenkins, A., Sanchez, A., Kilgus, S. P., & Eklund, K. (2021). Profiles of rater dis/agreement within universal screening in predicting distal

outcomes. *Journal of Psychopathology and Behavioral Assessment*, 1-14.

<https://doi.org/10.1007/s10862-021-09869-0>

von der Embse, N. P., Kilgus, S. P., Solomon, H. J., Bowler, M., & Curtiss, C. (2015). Initial development and factor structure of the Educator Test Stress Inventory. *Journal of Psychoeducational Assessment*, 33(3), 223-237. <http://dx.doi.org/10.1177/0734282914548329>.

von der Embse, N.P., Putwain, D. & \*Francis, G. (2021). Interpretation and Use of the Multidimensional Test Anxiety Scale. *School Psychology*, 36(2), 86–96.  
[doi.org/10.1037/spq0000427](https://doi.org/10.1037/spq0000427)

Zeidner, M. (1998). *Test anxiety: The state of the art*. New York, NY: 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* (pp. 141-163). New York, NY: Guilford Press.

Table 1  
Descriptive Statistics for MTAS Items

Variable	Mean	Variance	Skewness	Kurtosis
MTAS1	3.919	0.192	-1.005	-0.194
MTAS2	3.581	0.956	-0.655	0.695
MTAS3	3.961	0.962	-1.157	-0.225
MTAS4	2.650	0.935	0.291	1.086
MTAS5	3.870	1.578	-1.037	-1.068
MTAS6	3.597	0.936	-0.716	0.942
MTAS7	3.695	0.994	-0.805	-0.095
MTAS8	3.180	1.175	-0.168	-0.107
MTAS9	3.797	1.476	-0.929	-0.991
MTAS10	3.574	1.098	-0.595	0.381
MTAS11	3.582	1.027	-0.682	-0.324
MTAS12	2.803	1.350	0.179	-0.439
MTAS13	3.552	1.640	-0.606	-1.096
MTAS14	3.317	1.382	-0.231	-0.634
MTAS15	3.905	1.667	-1.152	-1.172
MTAS16	2.463	0.992	0.478	1.070
Worry	15.139	11.712	-0.845	0.195
Variable	Mean	Variance	Skewness	Kurtosis
Tension	15.143	11.722	-0.956	-0.025
Cog Inter	14.069	11.7	-0.535	0.586
Phys Ind	11.096	13.068	0.114	0.636
TA Total	55.447	1.336	-0.534	-0.581

Table 2  
Descriptive Fit Statistics for LPA Models

# of Profiles	Log Likelihood	# of Free Parameters	AIC	BIC	Sample-Size Adjusted BIC	LMR-LRT	Adjusted LMR-LRT	Bootstrap LRT	Entropy
1	-9931.710	8	19879.421	19917.998	19892.591	-	-	-	-
2	-9324.464	13	18674.927	18737.616	18696.329	1214.493 <i>p</i> =.000	1179.903 <i>p</i> =.000	1214.493 <i>p</i> =.000	0.867
3	-9032.983	18	18101.966	18188.766	18131.600	582.961 <i>p</i> =.000	566.358 <i>p</i> =.000	582.961 <i>p</i> =.000	0.824
4	-8939.507	23	17925.013	18035.924	17962.879	186.953 <i>p</i> =.0429	181.628 <i>p</i> =.0458	186.953 <i>p</i> =.0000	0.800
5	-8864.184	28	17784.367	17919.389	17830.464	150.646 <i>p</i> =.0023	146.355 <i>p</i> =.0026	150.646 <i>p</i> =.0000	0.821

Table 3  
 Mean Subscale Scores and Sample Percentages for Latent Profiles

	2 Profile Model						3 Profile Model						4 Profile Model						5 Profile Model					
	W	CI	T	PI	Total	% of sample	W	CI	T	PI	Total	% of sample	W	CI	T	PI	Total	% of sample	W	CI	T	PI	Total	% of sample
1	11	11	10	7	39	24	15	13	15	10	53	50	14	13	14	8	49	25	6	8	6	5	25	3
2	17	15	17	13	62	76	10	11	9	6	36	15	9	11	8	6	34	12	11	12	9	6	38	11
3							18	16	18	15	67	35	19	17	19	16	71	21	14	13	14	8	49	25
4													16	15	16	12	59	43	16	15	16	12	59	41
5													19	17	19	16	71	20	19	17	19	16	71	20

*\*Each of the subscales scores was out of a total score of 20 and each of the total scores were out of a total score of 80*