

Article

A General Math Anxiety Measure

Stephanie Kelly ^{1,*} , Stephen M. Croucher ², Kyung Yong Kim ³, Tatiana Permyakova ⁴ , Elira Turdubaeva ⁵, Kenneth T. Rocker ², Nadirabegim Eskiçorapçı ⁶, Gulzada Stanaliev ⁷, Bakyt Orunbekov ⁷ and Sucharat Rimkeeratikul ⁸ 

¹ Department of Business Information Systems and Analytics, North Carolina A&T State University, Greensboro, NC 27411, USA

² School of Communication, Journalism, and Marketing, Massey University, Wellington 6021, New Zealand; s.croucher@massey.ac.nz (S.M.C.); ktrock@ncat.edu (K.T.R.)

³ Department of Educational Research Methodology, University of North Carolina at Greensboro, Greensboro, NC 27412, USA; k_kim9@uncg.edu

⁴ School of Foreign Languages, HSE University, 614107 Perm, Russia; perm1@hotmail.com

⁵ Department of Media and Communications, University of Central Asia, Bishkek 720040, Kyrgyzstan; eliraturdubayeva@gmail.com

⁶ Public Relations and Publicity Department, Yeditepe University, Istanbul 34755, Turkey; nadira.eskicorapci@yeditepe.edu.tr

⁷ Department of Journalism, Kyrgyz-Turkish Manas University, Bishkek 720044, Kyrgyzstan; gulzada.stanaliyeva@manas.edu.kg (G.S.); bakytbek.orunbekov@manas.edu.kg (B.O.)

⁸ Language Institute, Thammasat University, Bangkok 10200, Thailand; sucharat.r@litu.tu.ac.th

* Correspondence: sekelly@ncat.edu; Tel.: +1-336-285-4903

Abstract: Math anxiety is a psychological burden that can hinder individuals across their lifetimes. However, the current literature lacks a valid measure of math anxiety that can be used across instructional modalities and among non-student populations. As such, it is difficult to assess math anxiety in virtual learning environments, track math anxiety across lifetimes, or determine the utility of math anxiety inoculations for non-student populations. This study presents a validity portfolio for a generalized math anxiety measure that can be used across teaching modalities, across lifetimes, and is simple enough to be used cross-culturally. The measure yielded evidence of validity when used in all tested samples: the United States (student and non-student samples), New Zealand (student and non-student samples), Kyrgyzstan (non-student sample), Turkey (non-student sample), Russia (non-student sample), and Thailand (non-student sample). The data support the use of the new math anxiety measure free of context.

Keywords: math anxiety; validity; measurement; cross-cultural



Citation: Kelly, S.; Croucher, S.M.; Kim, K.Y.; Permyakova, T.; Turdubaeva, E.; Rocker, K.T.; Eskiçorapçı, N.; Stanaliev, G.; Orunbekov, B.; Rimkeeratikul, S. A General Math Anxiety Measure. *Educ. Sci.* **2022**, *12*, 370. <https://doi.org/10.3390/educsci12060370>

Academic Editor: Molly M. Jameson

Received: 28 March 2022

Accepted: 16 May 2022

Published: 25 May 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

It is estimated that 93% of Americans, ranging across all age groups, suffer from some degree of math anxiety [1,2]. Math anxiety is a form of angst one experiences when working with numbers, but unlike the name implies, it is not constrained to traditional mathematics experiences [3]. Individuals may experience math anxiety in any quantitative reasoning activity, such as balancing a checkbook, reading a quantitative research study, interpreting election poll results, or thinking critically about whether an infographic accompanying televised news may be misleading.

In short, math anxiety, though studied primarily among student populations, is something that can affect individuals across their lifetimes. As prevalent as math anxiety is, scholars are limited in their ability to study it due to the measures available. The most popular math anxiety measures are the Programme for International Student Assessment (PISA) [4], the Abbreviated Math Anxiety Scale (AMAS) [5], and the Math Anxiety Rating Scale-Revised (MARS-R) [6]. These measures are all designed to assess math anxiety by specifically referencing physical math classrooms and math class preparation experiences.

As such, existing measures of math anxiety are not appropriate to assess math anxiety in virtual learning experiences or across a lifetime. These are both major limitations given the prevalence of virtual learning and the need to identify how well classroom inoculations against math anxiety hold once adults reach the workplace. As such, scholars lack the ability to assess the long-term impact of classroom interventions for math anxiety as well as math anxiety outside of the traditional classroom as it would be experienced in workplace trainings or during much of pandemic education.

Most research on adult math anxiety has primarily used the MARS-R for assessment, and the samples for these studies have primarily been adult students [2]. The MARS-R is the most appropriate existing measure for adults because it contains the fewest items among the three measures that necessitate thinking of resources and contexts that may exist only inside of a traditional math class, such as looking through the pages of a math text or physically walking to a mathematics classroom. Yet, the MARS-R has recently been identified as having several validity issues, which will be overviewed later in the paper [7–10]. As such, to better assess the longitudinal impact of math anxiety interventions across a lifetime, a new measure of math anxiety is needed that can be used inside and outside of the traditional classroom.

2. Math Anxiety and Measurement

Math anxiety is the unease some individuals experience when working with numbers [3,11]. Math anxiety symptoms can range from the psychological, such as feelings of dread and depression, to the physical, such as sweating hands and a racing heart [3,12]. Math anxiety is most problematic for students in that it is a barrier that prevents capable students from performing well in the math classroom.

The working memory, which is the brain's processing power in consciously active cerebral processes, is responsible for reasoning through steps or problems and for managing our behaviors and wellbeing moment to moment [9]. This means that the working memory is the resource students need to rely upon to solve a math problem, but also the resource needed to manage their anxiety symptoms [13,14]. As such, the more math anxiety symptoms a student has, the more working memory they must allot to managing those symptoms, leaving little left for solving the problems. So, although having math anxiety does not imply one is bad at or incapable of doing math, it does rob one of the working memory resources needed to be successful [9]. Therefore, many individuals who are perfectly capable of performing well at quantitative activities struggle not because they are bad at these activities, but because they do not have enough working memory left to focus on the activities once they have managed their math anxiety symptoms.

Math anxiety typically begins to develop at a young age when children observe parents or teachers exhibiting signs of math anxiety or phobia [15,16]. Children assume something severe or difficult enough to intimidate their trusted sources of knowledge must indeed be beyond comprehension, causing them to adopt their own apprehension towards mathematics. The earlier and more severe the symptoms of math anxiety when they arrive, the sooner students tend to begin disliking and avoiding mathematics. Beliefs about one's mathematical ability consistently relate to students' math anxiety [17]. A meta-analysis [18] found that math anxiety increases throughout the stages of education as learners accumulate more negative number-related experiences, typically peaking at high school and plateauing throughout adulthood.

Two concepts often discussed along with math anxiety in the literature are math self-concept and math self-efficacy. Math self-efficacy is one's belief that one can do math, and math self-concept is perceived self-worth based on one's ability to do math [19]. Both math self-efficacy and math self-concept are consistently found to be negatively related to math anxiety in the literature [19–21]. Like math anxiety, both math self-concept and math self-efficacy are predictors of math performance. Scholars [22] have found math self-efficacy to be a stronger predictor of a student's ability to do math than math self-concept. Math

self-efficacy is a stronger predictor of performance because math self-efficacy and social judgements create math self-concept [23].

Instructional communication research has identified that instructor immediate behaviors can serve as an intervention for math anxiety [8–10]. Instructional immediate behaviors are those behaviors that reduce perceived psychological distance in students. Nonverbal behaviors such as smiling and making eye contact served as an intervention for math anxiety in U.S. classrooms [9], and verbal behaviors such as demonstrating a moral code and attempting to build a rapport with students served as an intervention for math anxiety in the Chinese classroom [8]. Instructors' poor lecturing behaviors and aggressive communication with students increased their math anxiety [10].

Math anxiety has primarily been measured through the MARS-R [6] assessment, which was a reduction of the Math Anxiety Rating Scale (MARS). This is the most common measure both in classrooms and in adult non-student assessments [2]. Most studies using the MARS-R have not looked at its factor structure, relying on Hopko's results [6]. To be clear, this reliance on this work is not necessarily an example of neglect, as the 12-item assessment follows excellent practices of measurement development. Despite its era of development, the MARS-R cannot be condemned for any of the common measurement development errors scholars are identifying in older measures today. For example, a common pitfall of many older measures is the use of negatively worded items that were written with the intention of being reverse-coded [24,25]. These negatively worded items often created a second false factor, only visible through confirmatory factor analysis. The MARS-R measure was ahead of its time in avoiding this error, maintaining its unidimensional measurement structure across time [10]. However, recent studies that have examined the factor structure have found evidence of validity issues [8–10], indicating that though the factor structure has held, some of the items have become weak or otherwise problematic.

The reason the MARS-R items vary in their utility in modern use is simply that the nature of teaching mathematics and working with numbers has changed [9]. The MARS-R references events that were common in learning and practicing math in 2003 that are simply not common today. For example, three of the items reference using a physical book, rather than an e-book or Google to search for information, which does not reflect the actual practice of doing math today as it did in 2003. As the behaviors and references of people change over time, measures often lose their utility through no fault of the designer [26]. However, these measures become inappropriate to use, as they can no longer truly capture the construct through the invalidation of time. This is becoming the case for the MARS-R. With the prevalence and rapid evolution of technology in teaching, it is important to develop a measure that will not be dated as instructional delivery changes [24,27].

An additional limitation of using these specific math-related activities is that they do not allow participants to respond by considering the array of ways in which they interact with numbers. Many of the MARS-R items reference activities that would not occur outside of a math classroom. This is because the MARS-R was never intended to be used to assess math anxiety outside of a traditional classroom. This means that it cannot be used well in a workplace setting or for virtual learning, which many students experienced during the pandemic. Other math anxiety measures are similarly limited to the classroom. The AMAS math anxiety assessment was designed to assess math anxiety among students in face-to-face classrooms [5]. The PISA was developed along with math self-concept, math self-efficacy, and math interest assessments from a large-scale data collection from secondary students in Belgium [28]. This PISA math anxiety assessment was designed, however, to be used only for students currently enrolled in or preparing to enroll in a math classroom, as the items reflect current and future math classroom experiences. As such, it cannot be used to assess general math anxiety among an adult population or for students learning outside the traditional classroom.

Therefore, to address the limitations of current measures, a new assessment of math anxiety is needed. This assessment should meet the following goals:

1. The proposed measure can be used to assess math anxiety across contexts, which include varied learning modalities as well as classroom vs. workplace settings.
2. The proposed measure will not be bound to specific mathematics activities that may cause the measure to age out of utility.

Additionally, a major limitation to many bodies of emotive and perceptual research, such as math anxiety, is that there is no measure designed to be generalizable enough to be utilized cross-culturally [24]. Because of this, it is not possible to accurately assess how variables such as math anxiety and math anxiety intervention effects differ across classroom cultures. To address this gap in the literature, a third goal is set:

3. The measure should strive for simple wording that positions it to retain linguistic equivalence if used cross-culturally.

One of the reasons it is difficult to test social and psychological constructs cross-culturally is that the same measures cannot be used to assess the same construct when words take on different connotative meanings across cultures. The fewer words a measure contains, the fewer opportunities a measure has to lose linguistic equivalence [24]. Therefore, the measure type most likely to fit these needs is a semantic differential. Recurring themes in the definition of math anxiety include the terms *anxious*, *nervous*, *worry*, *panic*, and *uncertainty* [2,3,5–7,9,10]. After reviewing the math anxiety literature, the following items were developed to create the proposed math anxiety semantic differential measure:

Directions: *The following is a list of words that may describe how you feel when working with numbers. For each item, please select the points that indicate how you feel when working with numbers.*

1. *Anxious* _____ *Not Anxious*;
2. *Nervous* _____ *Not Nervous*;
3. *Worried* _____ *Not Worried*;
4. *Panicked* _____ *Not Panicked*;
5. *Uncertain* _____ *Certain*.

It is expected that if the measure yields evidence of validity, it will retain the unidimensional nature observed in other math anxiety measures and hold the same theoretical relations with other constructs [6,19]. As such, the following hypotheses are proposed:

Hypothesis 1 (H1). *Math anxiety is a unidimensional construct;*

Hypothesis 2 (H2). *The proposed math anxiety measure will relate negatively to math self-efficacy;*

Hypothesis 3 (H3). *The proposed math anxiety measure will relate negatively to math self-concept.*

It is also expected that the proposed measure will relate to the other measures of math anxiety. Because this measure assesses sentiments toward all quantitative reasoning activities rather than specific examples, the correlation may be moderate to strong. Therefore, the following hypothesis is proposed:

Hypothesis 4 (H4). *There will be a moderate to strong positive correlation between the new math anxiety measure and the MARS-R.*

Finally, if the wording of items is simple enough, the measure should retain factor structure well in international populations. Therefore, a final hypothesis is proposed:

Hypothesis 5 (H5). *The proposed math anxiety measure will retain its factor structure cross-culturally.*

3. Materials and Method

3.1. Participants

In total, 1,977 individuals participated in this study. Participants were located in the U.S. ($n = 548$), New Zealand ($n = 523$), Russia ($n = 204$), Kyrgyzstan ($n = 250$), Turkey ($n = 204$), and Thailand ($n = 248$). The first four hypotheses could be addressed with a U.S.-student sample, using the proposed measure and other U.S.-validated measures. The fifth hypothesis required non-U.S. participants. Data were collected from six countries in total. To test the fifth hypothesis, it was important to examine the function of the measure in at least one English-speaking country and at least one country that would require a language translation to be able to distinguish between potential validity threats due to cultural meaning vs. linguistic equivalence [24]. To collect the data, personal contacts of the lead researchers were consulted. This resulted in the identification of researchers willing to collect data in New Zealand, Russia, Kyrgyzstan, Turkey, and Thailand. Serendipitously, the inclusion of Russia, Kyrgyzstan, Turkey, and Thailand also helped redress the lack of research conducted on these regions [29], as noted by scholars who point out the limited representation of research samples from post-Soviet and West Asian cultures.

U.S. students. A total of 254 U.S. subjects participated in this study. Among this number, 114 were male and 140 were female. All subjects were college students. Their class rank broke down as follows: 6 freshmen, 111 sophomores, 104 juniors, and 33 seniors. Student major broke down as follows: 213 business, 8 STEM, 6 fine arts, 3 humanities, 4 social science, and 20 other. Subjects were on average 20.73 ($SD = 3.84$) years old.

U.S. adults. A total of 294 U.S. subjects participated in this study. Among this sample, 92 were males, 198 were females, 3 identified as other, and 1 chose not to identify sex. Education level broke down as follows: 2 lower than high school, 78 high school, 72 some university, 114 university degree or equivalent, 2 graduate education, 23 graduate degree, 1 postgraduate degree, and 2 who chose not to disclose education. Occupational fields broke down as follows: 23 technology, 22 food services, 25 education, 39 health, 38 sales, 29 specialized business, 20 skilled labor, 92 other, and 6 who chose not to disclose their occupation. Average age of subjects was 34.5 ($SD = 10.79$) years old.

New Zealand students. An additional student sample was solicited in New Zealand representing 210 students. Among the sample, 37 were male and 173 were female. Class rank broke down as follows: 45 freshmen, 29 sophomores, 18 juniors, and 118 seniors. Student major broke down as follows: 39 business, 72 STEM, 22 fine arts, 3 humanities, 16 social science, 15 education, 6 law, and 15 other. Subjects were on average 20.01 ($SD = 2.18$) years old.

New Zealand adults. A sample of 313 New Zealand adults was also collected. Among the sample, 168 subjects identified as male, 144 identified as female, and 1 identified as other. Education level broke down as follows: 22 high school, 75 some university, 175 university graduates, 17 some graduate education, 23 graduate degrees, and 1 postgraduate degree. Occupational fields broke down as follows: 26 technology, 33 sales, 33 health, 23 education, 46 skilled labor, 18 specialized business, 16 food services, 8 government, 12 hospitality/tourism, 90 other, and 8 subjects chose not to identify their occupation. Subjects' average age was 40.9 ($SD = 1.58$) years old.

Russia. The sample consisted of 204 Russian adults. Among these subjects, 70 were male, 114 were female, and 20 chose not to disclose their sex. Their education level broke down as follows: 1 GED, 32 high school graduates, 9 some college, 52 college graduates, and 110 professional or graduate education. Professional occupation area broke down as follows: 9 agriculture, 31 education, 13 engineering, 31 medical, 7 military, 39 skilled labor, 43 sales, 25 other, and 6 non-disclosed. The average age of subjects was 35.0 ($SD = 11.15$) years old.

Turkey. Among the 204 Turkish participants, the average age of subjects was 32.1 ($SD = 8.60$). Among subjects, 114 were male and 90 were female. Education broke down as follows: 6 some high school, 7 GED, 13 high school graduates, 13 with some college, 99 college graduates, and 48 with graduate or professional education. Occupation types

broke down as follows: 4 agriculture, 66 education, 25 engineering, 17 medical, 3 military, 12 skilled labor, 30 sales, and 47 other.

Kyrgyzstan. A total of 205 subjects participated from Kyrgyzstan. The average age of participants was 30.08 ($SD = 12.24$) years old. There were 70 males, 128 females, and 7 subjects who chose not to reveal their sex. Education of subjects broke down as follows: 127 with some college or vocational training, 28 with some high school education, 31 high school graduates, 17 college graduates, and 2 subjects who chose not to disclose their education level. Occupational fields broke down as follows: 3 agriculture, 52 education, 2 medical, 21 skilled labor, 7 sales, 5 technical, 107 other, and 8 who chose not to disclose their occupation.

Thailand. In total, 248 subjects participated from Thailand. The average age of participants was 26.73 ($SD = 8.04$). There were 76 males and 172 females. Education of subjects broke down as follows: 2 high school graduates, 112 some college, 74 college graduates, and 60 with graduate education. Occupation fields broke down as follows: 6 accounting, 139 education, 10 engineering, 26 military, 8 sales, and 59 other.

3.2. Procedure

U.S. students. Students were recruited to participate at a moderately sized southeastern university in the U.S. At this university, the business communication course is open to all majors but required for all business majors with degrees including management, entrepreneurship, economics, supply chain, marketing, business education, accounting, and finance. This course is designed with a research participation component, so students were presented with this study as one of the many they were eligible to participate in during the semester to earn credit. The study link was made available through the learning management system. Students who chose to participate were navigated from their learning management system to an online informed consent from which they could access the online questionnaire. Because a measure used to test hypothesis four was designed to be used in reference to a physical classroom, all participants acknowledged that they had had a face-to-face math class within the last two years before beginning the questionnaire. The average completion time for this questionnaire was 7 min.

U.S. adults, New Zealand adults, and New Zealand students. Data were collected in New Zealand and among U.S. adults via Qualtrics participant recruitment. Qualtrics made the online questionnaire link available to potential participants in exchange for approximately USD 1 in compensation. As with the U.S. sample, the average completion time for the questionnaire was approximately seven minutes.

Kyrgyzstan, Russia, Turkey, and Thailand. Before data were collected in Kyrgyzstan, Russia, Turkey, and Thailand, the questionnaires were first translated into Russian, Turkish, and Thai, respectively, by native speakers. The questionnaire was then back-translated by separate researchers into English to assure linguistic equivalence between the versions of the instrument. The reliability scores for the translations were 0.83 (Russian), 0.87 (Turkish), and 0.75 (Thai). Potential subjects were recruited via email and word of mouth. In Russia, Turkey, and Kyrgyzstan, respondents who were willing to complete the questionnaire were presented with a pencil and paper copy of the questionnaire to complete individually while the researcher waited on site. In Thailand, individuals willing to complete the questionnaire were emailed a link to an electronic version of the questionnaire hosted on Google Forms. The questionnaires were returned directly to the researchers with no compensation. Respondents needed approximately 10 min to complete the questionnaire.

3.3. Instrumentation

Math Anxiety. Math anxiety was assessed through two measures: the original measure described in the literature review and MARS-R measure. Descriptive statistics for all measures given to the U.S. sample are in Table 1.

Table 1. Descriptive statistics—finalized measures of U.S. student sample.

	Mean	SD	Min–Max	Skewness	Kurtosis	α
Math Anxiety New	4.70	1.60	1.00–7.00	−0.29	−0.65	0.91
Math Anxiety Old	3.80	1.66	1.00–7.00	0.11	−0.92	0.92
Math Self-Efficacy	4.77	1.39	1.00–7.00	−0.53	−0.23	0.89
Math Self-Concept	3.88	1.13	1.00–7.00	0.13	−0.48	0.83

Math Self-Efficacy. Math self-efficacy was measured through the PISA measure [4]. This measure consists of seven semantic differential items with seven-point response scales ranging from *Disagree Strongly* to *Agree Strongly*.

Math Self-Concept. Math self-concept was measured through the PISA measure [4]. This measure consists of five semantic differential items with seven-point response scales ranging from *Disagree Strongly* to *Agree Strongly*.

4. Results

The initial validity tests for the proposed measure had to be conducted in a U.S. student sample so the MARS-R measure could be used in its intended population for concurrent validity checks. This was necessary because all previously validated measures that have an established record of relating to math anxiety, those needed to complete a test of parallelism in developing a validity portfolio, have been validated only among student samples. Therefore, U.S. student sample data were used to test the first four hypotheses.

4.1. Content Validity

The first step in testing the content validity (i.e., fit of proposed and observed factor structure) of the measure was to run an exploratory factor analysis (EFA) on the data to observe the factor structure of the measure when unconstrained. The EFA was run with SPSS using a principal component method. All items were loaded on a single factor with a magnitude of 0.86 or higher. The factor loadings are shown in Table 2. Therefore, the EFA supported the unidimensional factor structure of the proposed measure.

Table 2. Exploratory factor analysis—U.S. students.

Item	Loading Component 1
Anxious	0.86
Nervous	0.92
Panicked	0.89
Uncertain	0.86
Worried	0.93

The second step in assessing the validity of the proposed measure was a confirmatory factor analysis (CFA) to confirm the factor structure. CFA has two phases: (1) a test of internal consistency and (2) a test of parallelism. The test of internal consistency ascertains that there are no problematic or weak items in the measure. The test of parallelism ascertains that the proposed measure assesses only the construct it is purported to measure. The Mplus weighted least square mean and variance adjusted algorithm was used to complete the CFA. For categorical variable measurement such as adopted for this proposed measure, there is debate regarding the most appropriate fit cutoffs [30]. However, it is generally accepted that the fit is acceptable if at least three of the five conditions are met [20]: Tucker–Lewis index (TLI) ≥ 0.95 , comparative fit index (CFI) ≥ 0.95 , standard root mean residual (SRMR) ≤ 0.08 , root mean residual error approximation (RMSEA) ≤ 0.05 , and a non-significant chi-square test. Notably, RMSEA is based on chi-square analyses, so both RMSEA and chi-square are extremely sensitive to minor misfit; as such, a measure is still considered to have ample evidence of content validity if RMSEA and chi-square alone are poor [31,32].

The fit statistics for the proposed five-item math anxiety measure from the test of internal consistency were as follows: TLI = 1.00, CFI = 1.00, RMSEA = 0.14, SRMR = 0.09, and $\chi^2(10, N = 254) = 7012.04, p < 0.001$. The TLI and CFI were well within the boundaries of good fit, but the RMSEA and SRMR were not, and the chi-square was statistically significant. Given this evidence of misfit, the standard residual covariance matrix was examined. One item (*worried, not worried*) was observed to cause a statistically significant amount of residual error across multiple other items. After this item was dropped, TLI, CFI, and SRMR fell within acceptable range: TLI = 0.99, CFI = 0.98, RMSEA = 0.20, SRMR = 0.01, and $\chi^2(2, N = 254) = 21.93, p < 0.001$. Therefore, the item (*worried, not worried*) was dropped from subsequent analyses. The measurement model with standardized regression weights is shown in Figure 1.

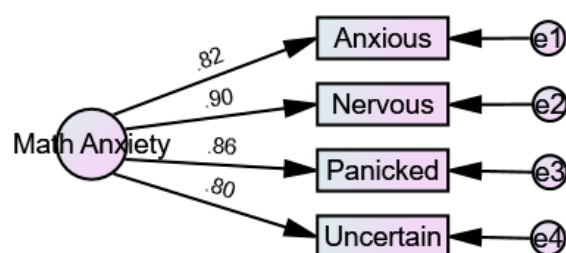


Figure 1. Measurement model with standardized regression weights.

Next, tests of internal consistency were run on both the math self-concept and math self-efficacy measures. The fit statistics for the math self-efficacy measure were as follows: TLI = 0.80, CFI = 0.87, RMSEA = 0.33, SRMR = 0.09, and $\chi^2(21, N = 254) = 2870.12, p < 0.001$. All fit statistics were very poor. Three items (*Solving an equation like $3x+5=17$* , *Finding the actual distance between two places on a map with a 1:10,000 scale*, and *Calculating the petrol consumption rate of a car*) were found to cause a statistically significant amount of residual error across multiple other items. As such, they were dropped before the test of parallelism, yielding the following fit statistics: TLI = 0.98, CFI = 0.99, RMSEA = 0.20, SRMR = 0.01, and $\chi^2(6, N = 254) = 3620.19, p < 0.001$. Although the RMSEA was still elevated and the chi-square test was statistically significant, the other fit statistics were good, meaning the respecified measures were appropriate for use in future tests [31,32]. The math self-concept measure was likewise examined, yielding fit statistics as follows: TLI = 0.98, CFI = 0.99, RMSEA = 0.20, SRMR = 0.01, and $\chi^2(10, N = 254) = 6160.03, p < 0.001$. With acceptable TLI, CFI, and SRMR, the original measure is appropriate for use in hypothesis testing. When the measures were compared in the test of parallelism, no problematic items were identified, indicating that the proposed math anxiety measurement items were not measuring math self-efficacy or math-self-concept by accident (TLI = 0.95, CFI = 0.96, RMSEA = 0.07, SRMR = 0.04, and $\chi^2(62, N = 254) = 147.27, p < 0.001$).

4.2. Concurrent Validity

Concurrent validity evidence is provided if a proposed measure is observed to have the expected relationship with another measure as predicted by theory or past research. Therefore, to test concurrent validity, hypothesis two predicted a negative relationship between the new math anxiety measure and math self-concept. This relationship was supported ($r = -0.53, p < 0.00$). The third hypothesis predicted a negative relationship between the new math anxiety measure and math self-efficacy. This relationship was also supported ($r = -0.39, p < 0.001$). Therefore, the proposed math anxiety measure yielded evidence of concurrent validity. The correlation matrix for these measures is shown in Table 3.

Table 3. Correlation matrix—U.S. student sample.

	1	2	3
1. Math Anxiety New			
2. Math Anxiety Old	0.54 *		
3. Math Self-Efficacy	−0.39 *	−0.34 *	
4. Math Self-Concept	−0.53 *	−0.58 *	0.50 *

* $p < 0.05$.

4.3. Convergent Validity

A measure is said to have evidence of convergent validity if two measures of the same construct are moderately to strongly correlated. Therefore, the fourth hypothesis predicted a positive relationship between the proposed math anxiety measure and the MARS-R. Before testing this hypothesis, the MARS-R was subjected to CFA. The fit statistics were acceptable within this dataset: TLI = 0.92, CFI = 0.94, RMSEA = 0.21, SRMR = 0.06, and $\chi^2(66, N = 254) = 8956.77, p < 0.001$. The correlation between the two math anxiety measures was $r = 0.54$ ($p < 0.001$). Although this is a moderate correlation, it provides evidence of convergent validity, as it compares measures of general and classroom-specific math anxiety.

5. U.S. Adult Sample

Next, the measure was utilized with a non-student U.S. adult population to examine the factor structure of the measure in a non-student sample from the same culture as the student sample. If the measure is indeed valid for students and non-students, then the four-item unidimensional measure should have acceptable fit statistics among the U.S. adult sample as well. The fit statistics were as follows: TLI = 0.99, CFI = 1.00, RMSEA = 0.16, SRMR = 0.01, and $\chi^2(2, N = 294) = 16.64, p < 0.001$. The TLI, CFI, and SRMR fell within the range of acceptable fit. Again, chi-square tests and RMSEAs (which are calculated based upon the chi-square test) are extremely sensitive to even minor misfit, so measures are considered to be of acceptable, but not exemplary fit, if CFI, GFI, and SRMR are strong but the RMSEA and/or chi-square test alone are problematic [32,33]. Therefore, this measure has adequate evidence of content validity within the U.S. adult non-student sample.

6. International Samples

The final hypothesis predicted that the measure would retain its fit in international samples. The measure was disseminated in New Zealand (both student and adult populations), Russia, Kyrgyzstan, Turkey, and Thailand. Table 4 shows the fit statistics and Table 5 shows the descriptive statistics for the measure within these samples. The TLI, CFI, RMSEA, and SRMR were all within the range of good fit, and the chi-square test was not statistically significant for New Zealand (both student and adult samples) and Kyrgyzstan. For the Turkey sample, TLI, CFI, and SRMR fell within acceptable range, and the chi-square test was statistically significant, but the RMSEA was elevated. For Russia and Thailand, the TLI, CFI, and SRMR fell within the range of good fit, but the RMSEA was elevated, and the chi-square was not statistically significant. Therefore, the measure shows excellent fit in the U.S., New Zealand, and Kyrgyzstani samples and acceptable fit in the Russian, Turkish, and Thai samples.

Table 4. Fit statistics—international samples.

	TLI	CFI	RMSEA	SRMR	χ^2
New Zealand Adult	1.00	1.00	0.02	0.01	$\chi^2(2, N = 313) = 3.15, p = 0.21$
New Zealand Student	1.00	1.00	0.00	0.00	$\chi^2(2, N = 210) = 0.33, p = 0.85$
Russia	0.99	1.00	0.19	0.01	$\chi^2(2, N = 204) = 16.94, p < 0.001$
Thailand	0.96	0.99	0.25	0.02	$\chi^2(2, N = 248) = 32.06, p < 0.001$
Turkey	0.99	1.00	0.10	0.02	$\chi^2(2, N = 204) = 2.98, p = 0.05$
Kyrgyzstan	1.00	1.00	0.00	0.00	$\chi^2(2, N = 205) = 0.53, p = 0.77$

Table 5. Descriptive statistics for non-U.S.-student samples.

	Mean	SD	Min–Max	Skewness	Kurtosis	α
U.S. Adult	4.79	1.80	1.00–7.00	−0.35	−0.90	0.92
New Zealand Adult	5.16	1.58	1.00–7.00	−0.57	−0.39	0.94
New Zealand Student	4.21	1.57	1.00–7.00	0.11	−0.87	0.91
Russia	5.36	1.64	1.00–7.00	−0.90	0.03	0.93
Thailand	4.23	1.47	1.00–7.00	−0.03	0.16	0.89
Turkey	4.85	1.51	1.00–7.00	−0.56	−0.36	0.85
Kyrgyzstan	5.17	1.41	1.00–7.00	−0.73	−0.01	0.72

To further examine the behavior of the measure across samples, a series of multi-group CFAs were ran using the U.S. student sample as the basis for comparison. Each measure pairing was tested for configural invariance (whether factor structures match across groups), metric invariance (whether factor loadings match across groups), and scalar invariance (whether thresholds are the same across groups). The results of these analyses are displayed in Table 6. As shown also with the unidimensional measurement tests, the factor structures were maintained. All groups except for Kyrgyzstan are metric invariant with the U.S. student sample, meaning the factor loadings maintained similar patterns of strength across groups. The results of the scalar invariance test indicated that although the measure maintained factor structure, different cultures seem to have different categorical thresholds for what constitutes “high” math anxiety. This is not a threat to validity, but something for math anxiety researchers to be aware of, as it may affect their interpretation of future research results.

Table 6. Fit comparison to U.S. student sample.

		TLI	CFI	RMSEA	SRMR	χ^2	$\Delta \chi^2$
U.S. Adult	Configural	1.00	1.00	0.18	0.01	$\chi^2(4, N = 294) = 38.89, p < 0.001$	
	Metric	1.00	1.00	0.12	0.01	$\chi^2(7, N = 294) = 35.80, p < 0.001$	$\chi^2(3, N = 294) = 6.55, p = 0.09$
	Scalar	1.00	1.00	0.07	0.02	$\chi^2(26, N = 294) = 56.99, p < 0.001$	$\chi^2(19, N = 294) = 35.30, p = 0.01$
New Zealand Adult	Configural	0.99	1.00	0.14	0.01	$\chi^2(4, N = 313) = 27.20, p < 0.001$	
	Metric	1.00	1.00	0.10	0.01	$\chi^2(7, N = 313) = 25.50, p < 0.001$	$\chi^2(3, N = 313) = 4.20, p = 0.24$
	Scalar	1.00	1.00	0.07	0.02	$\chi^2(29, N = 313) = 59.18, p < 0.001$	$\chi^2(19, N = 313) = 40.85, p < 0.001$
New Zealand Student	Configural	0.99	1.00	0.14	0.14	$\chi^2(4, N = 210) = 20.98, p < 0.001$	
	Metric	1.00	1.00	0.08	0.08	$\chi^2(7, N = 210) = 18.51, p = 0.01$	$\chi^2(3, N = 210) = 2.86, p = 0.41$
	Scalar	1.00	1.00	0.07	0.07	$\chi^2(26, N = 210) = 54.14, p < 0.001$	$\chi^2(19, N = 210) = 39.98, p < 0.001$

Table 6. Cont.

		TLI	CFI	RMSEA	SRMR	χ^2	$\Delta \chi^2$
Russia	Configural	0.99	1.00	0.19	0.01	$\chi^2(4, N = 204) = 35.89$, $p < 0.001$	
	Metric	0.99	1.00	0.14	0.01	$\chi^2(7, N = 204) = 39.02$, $p < 0.001$	$\chi^2(3, N = 204) = 5.48$, $p < 0.001$
	Scalar	1.00	1.00	0.08	0.02	$\chi^2(26, N = 204) = 31.56$, $p < 0.001$	$\chi^2(19, N = 204) = 31.56$, $p = 0.04$
Thailand	Configural	0.98	0.99	0.22	0.01	$\chi^2(4, N = 248) = 54.20$, $p < 0.001$	
	Metric	0.99	0.99	0.15	0.01	$\chi^2(7, N = 248) = 44.42$, $p < 0.001$	$\chi^2(3, N = 248) = 3.92$, $p < 0.001$
	Scalar	0.99	0.98	0.14	0.02	$\chi^2(26, N = 248) = 153.70$, $p < 0.001$	$\chi^2(19, N = 248) = 120.78$, $p < 0.001$
Turkey	Configural	0.99	1.00	0.15	0.01	$\chi^2(4, N = 204) = 24.013$, $p < 0.001$	
	Metric	1.00	1.00	0.10	0.01	$\chi^2(7, N = 204) = 21.90$, $p = 0.003$	$\chi^2(3, N = 204) = 2.84$, $p = 0.42$
	Scalar	0.99	0.99	0.11	0.03	$\chi^2(26, N = 204) = 99.52$, $p < 0.001$	$\chi^2(19, N = 204) = 81.57$, $p < 0.001$
Kyrgyzstan	Configural	0.99	1.00	0.11	0.01	$\chi^2(4, N = 205) = 15.33$, $p = 0.004$	
	Metric	0.99	1.00	0.12	0.01	$\chi^2(7, N = 205) = 31.78$, $p < 0.001$	$\chi^2(3, N = 205) = 16.07$, $p = 0.001$
	Scalar	0.99	1.98	0.12	0.04	$\chi^2(26, N = 205) = 105.62$, $p < 0.001$	$\chi^2(19, N = 205) = 89.93$, $p < 0.001$

7. Conclusions

Overall, these data provide validation evidence that supports the proposed math anxiety measure. The finalized measure and instructions are listed in Appendix A. This measure improves upon prior measures of math anxiety because it can be used across lifetimes, learning modalities, and contexts of mathematical reasoning. As such, this new generalized math anxiety measure can assess math anxiety regardless of whether students are learning in a traditional brick-and-mortar classroom or learning fully online; whether students are using digital learning materials or hands-on tools; or whether students are in a traditional mathematics course, statistics course, or any other course that utilizes numbers (e.g., accounting, economics, finance). This generalized math anxiety measure will also assess math anxiety outside of the classroom setting, so that math anxiety can be studied across lifetimes, allowing mathematics educators to assess how classroom interventions prepare adults to deal with math anxiety after they have completed school.

The original measurement model proposed five items, but four items were confirmed as part of the factor structure among the U.S. student sample. The item *worried, not worried* was lost through CFA and therefore not tested among the non-U.S. samples or the U.S. adult sample. This item was likely problematic because it was the only item of the five that implied some consideration of the future, whereas the other four could be answered by only thinking of dealing with numbers in the present moment. As such, hypothesis one, which predicted a unidimensional measurement model, was confirmed with the four supported items.

Hypotheses two, three, and four predicted positive relationships between the proposed math anxiety measure as well as the measures of math self-efficacy, math self-concept, and the MARS-R. The data were consistent with all three hypotheses. That statistically significant positive correlations were observed between the new generalized math anxiety model and constructs that it should theoretically be positively related to provides evidence of convergent validity. Further, the moderate positive correlation between the MARS-R and the new generalized math anxiety measure provided evidence of concurrent validity.

The data also supported, in agreement with hypothesis five, the conclusion that the four-item measure retained its factor structure in non-U.S. cultures. The measure had excellent global fit statistics among the New Zealand, Kyrgyzstan, and Turkish cultures, and

acceptable global fit among the Russian and Thai samples. Though the data indicate that this measure likely adapts well across cultures, researchers are still encouraged to confirm the factor structure before use in U.S. and non-U.S. samples. While the psychological construct of math anxiety likely holds across cultures, lexical choices and translation failure could obscure the meaning of the measure across cultures or in the future as connotative meaning of the words used to build the items changes across age groups [24,26]. Researchers should never assume the validity of the measure and should perform their due diligence to confirm the factor structures to ensure sound measurement practices and sound inferences of data in future research [27,29]. Failing to do so threatens the integrity of science, preventing changes in human behavior from being detected and obscuring knowledge about the relationships between variables [27,34].

This study was not without limitations. First, it was not possible to collect an adult and student sample from each country. Second, the non-random sampling techniques resulted in bias in some of the samples. Most notably, the U.S. student sample represented mostly business majors, the Kyrgyzstan data represented mostly “other” occupations, and the Thai sample was highly educated. The same sampling bias was not present in any two of the samples utilized for this study, so the authors are not concerned that the measure is not representative, but random sampling would have made the study stronger.

In short, this measure advances the study of math anxiety by providing a valid measure that can be used for longitudinal studies across lifetimes and cross-cultural studies. The quantitative reasoning situations that induce math anxiety are encountered in all classroom types, across occupations, and throughout the lifetime whether one is a student taking an asynchronous online quantitative research methods course or an adult trying to understand election poll predictions. Now that this measure exists, scholars are encouraged to track math anxiety intervention research longitudinally across lifetimes, rather than limiting that research to populations engaged in formal, in-person mathematics education.

Author Contributions: Conceptualization, S.K. and S.M.C.; methodology, S.K., formal analysis, S.K. and K.Y.K.; data curation S.K., S.M.C., T.P., E.T., N.E., G.S., B.O. and S.R.; writing—original draft preparation, S.K. and K.T.R.; writing—review and editing, S.M.C. and K.Y.K.; All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and approved by the Institutional Review Board. Data collection in New Zealand, Turkey, Thailand, Russia and US (adult population) used IRB 4000024506 from Massey University. Data collection of US students used IRB 17-0152 from North Carolina A&T State University.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: IRB guidelines prevent the data from being published, but they can be reviewed individually by request by contacting the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Directions: *The following is a list of words that may describe how you feel when working with numbers. For each item, please select the points that indicate how you feel when working with numbers.*

1. Not Anxious _____ Anxious;
2. Not Nervous _____ Nervous;
3. Not Panicked _____ Panicked;
4. Certain _____ Uncertain.

References

- Blazer, C. Strategies for Reducing Math Anxiety. In *Information Capsule Research Services*; Miami-Dade County Public Schools: Miami, FL, USA, 2011. Available online: <https://files.eric.ed.gov/fulltext/ED536509.pdf> (accessed on 17 May 2022).
- Luttenberger, S.; Wimmer, S.; Paechter, M. Spotlight on Math Anxiety. *Psychol. Res. Behav. Manag.* **2018**, *11*, 311–322. [CrossRef] [PubMed]
- Ma, X. A Meta-Analysis of the Relationship Between Anxiety Toward Mathematics and Achievement in Mathematics. *J. Res. Math. Educ.* **1999**, *30*, 520–540. [CrossRef]
- Organisation for Economic Co-Operation and Development (OECD). *PISA 2003 Data Analysis Manual: SPSS Edition*; Organisation for Economic Co-Operation and Development (OECD): Paris, France, 2005.
- Hopko, D.R.; Mahadevan, R.; Bare, R.L.; Hunt, M.K. The Abbreviated Math Anxiety Scale (AMAS) Construction, Validity, and Reliability. *Assessment* **2003**, *10*, 178–182. [CrossRef] [PubMed]
- Hopko, D.R. Confirmatory Factor Analysis of The Math Anxiety Rating Scale-Revised. *Educ. Psychol. Meas.* **2003**, *63*, 336–351. [CrossRef]
- Johnson, L.L.; Kelly, S. Student Predispositions as Predictors Of Dissent Behaviors in Supply Chain Courses. *Decis. Sci. J. Innov. Educ.* **2020**, *18*, 270–290. [CrossRef]
- Kelly, S.; Liu, L.; Denton, Z.; Lee, C.; Croucher, S. Instructional Immediacy in the Chinese Quantitative Reasoning Classroom. *Sch. Sci. Math.* **2018**, *118*, 104–112. [CrossRef]
- Kelly, S.; Rice, C.; Wyatt, B.; Ducking, J.; Denton, D. Teacher Immediacy and Decreased Student Quantitative Reasoning Anxiety: The Mediating Effect of Perception. *Commun. Educ.* **2015**, *64*, 171–186. [CrossRef]
- Kelly, S.; Romero, A.; Morrow, J.A.; Denton, Z.; Ducking, J. Instructor misbehaviors and math anxiety. *Commun. Rep.* **2020**, *33*, 27–40. [CrossRef]
- Liu, F. Anxiety Towards Teaching Mathematics and Science: Correlation, Prevalence, and Intensity. *J. Math. Educ.* **2016**, *9*, 29–46. Available online: http://educationforatoz.net/images/Fuchang_Liu_2016.pdf (accessed on 13 May 2022).
- Plaisance, D.V. A Teacher's Quick Guide to Understanding Mathematics Anxiety. *La. Assoc. Teach. Math. J.* **2009**, *6*, 1–8. Available online: http://www.lamath.org/journal/vol6no1/anxiety_guide.pdf (accessed on 13 May 2022).
- Morsanyi, K.; Busdraghi, C.; Primi, C. Mathematical Anxiety is Linked to Reduced Cognitive Reflection: A Potential Road From Discomfort in the Mathematics Classroom to Susceptibility to Biases. *Behav. Brain Funct.* **2014**, *10*, 31. [CrossRef] [PubMed]
- Shaikh, S.N. Mathematics Anxiety Factors and Their Influence on Performance in Mathematics in Selected International Schools in Bangkok. *J. Educ. Vocat. Res.* **2013**, *4*, 77–85. [CrossRef]
- Beilock, S.L.; Willingham, D.T. Math Anxiety: Can Teachers Help Students Reduce It? Ask the Cognitive Scientist. *Am. Educ.* **2014**, *38*, 28–32. Available online: <https://dbweb01.aft.org/sites/default/files/periodicals/beilock.pdf> (accessed on 17 May 2022).
- Soni, A.; Kumari, S. The role of Parental Math Anxiety and Math Attitude in Their Children's Math Achievement. *Int. J. Sci. Math. Educ.* **2017**, *15*, 331–347. [CrossRef]
- Op't Eynde, P.; De Corte, E.; Verschaffel, L. Framing Students' Mathematics Related Beliefs. In *Beliefs: A Hidden Variable in Mathematics Education*; Leder, G.C., Pehkonen, E., Törner, G., Eds.; Kluwer: Alphen aan den Rijn, The Netherlands, 2002; pp. 13–37.
- Hembree, R. The Nature, Effects, and Relief of Mathematics Anxiety. *J. Res. Math. Educ.* **1990**, *21*, 33–46. [CrossRef]
- Jameson, M.M.; Fusco, B.R. Math Anxiety, Math Self-Concept, And Math Self-Efficacy in Adult Learners Compared to Traditional Undergraduate Students. *Adult Educ. Q.* **2014**, *64*, 306–322. [CrossRef]
- Ahmed, W.; Minnaert, A.; Kuyper, H.; van der Werf, G. Reciprocal Relationships Between Math Self-Concept and Math Anxiety. *Learn. Individ. Differ.* **2012**, *22*, 385–389. [CrossRef]
- Lee, J. Universals and Specifics of Math Self-Concept, Math Self-Efficacy, and Math Anxiety Across 41 PISA 2003 Participating Countries. *Learn. Individ. Differ.* **2009**, *19*, 355–365. [CrossRef]
- Pajares, F.; Miller, M.D. Role of Self-Efficacy and Self-Concept Beliefs in Mathematical Problem Solving: A path analysis. *J. Educ. Psychol.* **1994**, *86*, 193–203. [CrossRef]
- Bong, M.; Skaalvik, E.M. Academic Self-Concept and Self-Efficacy: How Different Are They Really? *Educ. Psychol. Rev.* **2003**, *15*, 1–40. [CrossRef]
- Croucher, S.M.; Kelly, S. Measurement Issues in Intercultural Communication. In *Communication Research Measures III: A Sourcebook*; Mazer, J., Graham, E., Eds.; Sage: Thousand Oaks, CA, USA, 2019; pp. 141–159.
- Kotowski, M.R.; Levine, T.R.; Baker, C.R.; Bolt, J.M. A Multitrait–Multimethod Validity Assessment of the Verbal Aggressiveness and Argumentativeness Scales. *Commun. Monogr.* **2009**, *76*, 443–462. [CrossRef]
- Autman, H.; Kelly, S. Reexamining the Writing Apprehension Measure. *Bus. Prof. Commun. Q.* **2017**, *80*, 516–529. [CrossRef]
- Kelly, S.; Westerman, D. Doing Communication Science: Thoughts on Making More Valid Claims. *Ann. Int. Commun. Assoc.* **2020**, *44*, 177–184. [CrossRef]
- Schulz, W.H. Mathematics Self-Efficacy and Student Expectations: Results from PISA 2003. In *Proceedings of the Annual Meetings of the 2005 American Educational Research Association*, Montreal, QC, Canada, 11–15 April 2005.
- Croucher, S.M.; Kelly, S.E.; Rahmani, D.; Burkey, M.; Subanaliev, T.; Galy-Badenas, F.; Lando, A.L.; Chibita, M.; Nyiransabimana, V.; Turdubaeva, E.; et al. A Multi-National Validity Analysis of the Self-Perceived Communication Competence Scale. *J. Int. Intercult. Commun.* **2020**, *13*, 1–12. [CrossRef]

-
30. Bowen, N.K.; Masa, R.D. Conducting measurement invariance tests with ordinal data: A guide for social work researchers. *J. Soc. Soc. Work Res.* **2015**, *6*, 229–249. [[CrossRef](#)]
 31. Kline, R.B. *Principles and Practice of Structural Equation Modeling*, 4th ed.; The Guilford Press: New York, NY, USA, 2016.
 32. Chen, F.; Curran, P.J.; Bollen, K.A.; Kirby, J.; Paxton, P. An Empirical Evaluation of the Use Of Fixed Cutoff Points in RMSEA Test Statistic in Structural Equation Models. *Sociol. Methods Res.* **2008**, *36*, 462–494. [[CrossRef](#)] [[PubMed](#)]
 33. Crede, M.; Harms, P. Questionable Research Practices When Using Confirmatory Factor Analysis. *J. Manag. Psychol.* **2019**, *34*, 18–30. [[CrossRef](#)]
 34. McEwan, B.; Carpenter, C.J.; Westerman, D.K. On Replication in Communication Science. *Commun. Stud.* **2018**, *69*, 235–241. [[CrossRef](#)]