Validation of the Math Anxiety Scale with the Rasch Measurement Model

Article *in* Mathematics Education Research Journal · June 2018

DOI: 10.1007/s13394-018-0244-8

2 authors:

Ibrahim Burak Olmez
MEF University

Safiye Bahar Ölmez
Kanuni Sultan Süleyman Training and Research Hospital Istanbul

Math Ed Res J (2019) 31:89–106
https://doi.org/10.1007/s13394-018-0244-8

Validation of the Math Anxiety Scale with the Rasch Measurement Model

İbrahim Burak Ölmez[1] **· Safiye Bahar Ölmez**[2]

Received: 25 December 2017 / Revised: 13 May 2018 / Accepted: 21 May 2018 /
Published online: 13 June 2018

Abstract The purpose of this study was to investigate the psychometric characteristics of the Math Anxiety Scale (MANX; Erol 1989, Unpublished master thesis, Bogazici University) with data collected from 952 middle school students in Turkey. The Rasch Rating Scale model was used to examine the MANX at the item level. The results revealed that although the MANX was sensitive to detect students with moderate levels of math anxiety and it was not targeted to identify those with very high and low math anxiety levels, it had high reliability and validity. Moreover, the majority of the MANX items were of good quality. The results of this study provide strong evidence for the validation of the MANX despite the need for deletion of eight misfit items and three items with the same item difficulties. Future research should consider possible revision or development of new items to capture gradations of challenges at the very high and low ends of the continuum.

Keywords Math anxiety · Rasch rating scale model · Item functioning · Middle school students

Introduction

Despite the importance attached to mathematics by policy-makers and educators around the world, proficiency in mathematics continues to be a persistent challenge for students and teachers in school settings. One major barrier to mathematical proficiency has consistently been considered as math anxiety, which can be defined as

✉ İbrahim Burak Ölmez
i.burakolmez@hotmail.com

Safiye Bahar Ölmez
safiyebaharolmez@gmail.com

[1] Department of Mathematics and Science Education, University of Georgia, Athens, GA 30602, USA

[2] Department of Psychiatry, School of Medicine, Duzce University, Duzce 81620, Turkey

"feelings of tension and anxiety that interfere with the manipulation of numbers and the solving of mathematical problems in a wide variety of ordinary life and academic situations" (Richardson and Suinn 1972, p. 551). Math anxiety has been consistently linked to several negative outcomes such as low academic performance (Ashcraft 2002), reduced working memory functioning (Ashcraft and Moore 2009; Young et al. 2012), low perceptions of one's own abilities (Hembree 1990), and deficits in solving mathematical problems (Ho et al. 2000). This, in turn, can cause many students to avoid taking high school and college mathematics courses and pursuing career paths that do not involve mathematics (Ma 1999).

In response to the need to identify math anxiety levels, researchers have developed self-report scales. The most popular scale for measuring math anxiety is the Mathematics Anxiety Rating Scale (MARS) developed by Richardson and Suinn (1972), and it has 98 items on a 5-point Likert scale. Because the MARS is a long scale to administer and is only suitable for adults, several versions of the MARS have been designed such as the Revised Mathematics Anxiety Rating Scale (RMARS; Alexander and Martray 1989), the Mathematics Anxiety Rating Scale for Adolescents (MARS-A; Suinn and Edwards 1982), and the Mathematics Anxiety Rating Scale for Elementary School Students (MARS-E; Suinn et al. 1988). An alternative scale to the MARS is the Mathematics Anxiety Scale (MAS) with 12 items on a 5-point Likert scale; it is appropriate for middle school and high school students (Fennema and Sherman 1976). The third scale, which is relatively less popular than the various forms of the MARS and the MAS, is the Mathematics Anxiety Questionnaire (MAQ) developed by Wigfield and Meece (1988). The MAQ has 11 items on a 7-point Likert scale and appropriate for use with middle school and high school students.

In addition to the Turkish translations of these three scales, there also exist a few scales originally developed in Turkish and have been used with Turkish population, including the Math Anxiety Scale (Bindak 2005) and the Math Anxiety Scale (MANX; Erol 1989). The Math Anxiety Scale developed by Bindak (2005) consists of 10 items with five response categories ranging from "always" to "never" and appropriate for use with elementary school students. On the other hand, the MANX, developed by Erol (1989), has 45 items with four response categories for each item ranging from "never" to "always." For the present study, we decided to use the MANX because it is the most widely employed math anxiety scale in Turkey. Several studies have administered the MANX to measure the math anxiety levels of middle grades students (e.g., Erden and Akgül 2010), of high school students (e.g., Ader and Erktin 2010), and of pre-service teachers (e.g., Aydin et al. 2009; Bekdemir 2010). The only study, however, that examined the psychometric characteristics of the MANX was the one conducted by Erktin et al. (2006), which used factor analysis to determine the underlying factors of the scale on a sample of 754 middle school and high school students. Erktin et al. identified four factors: (a) *test and evaluation anxiety*, which captures students' anxiety about exams and evaluation; (b) *apprehension of math lessons*, which captures students' anxiety related to mathematics lessons; (c) *use of mathematics in daily life*, which captures students' anxiety in daily situations that involve mathematics; and (d) *self-efficacy for mathematics*, which captures students' self-perceptions of themselves in mathematics. Given the factors revealed by Erktin et al., math anxiety can be considered as a unidimensional but complex construct where a higher level of math anxiety is characterized by high anxiety about exams and being evaluated by someone else, high anxiety about mathematics lessons, high anxiety about experiencing daily-life situations related to mathematics, and low self-efficacy involving mathematics, and vice versa.

Regarding the negative long-term impacts of math anxiety, it is important to have validated scales that fully capture the math anxiety levels of students for identifying those students and providing specific interventions based on the needs of each student. To achieve this aim, the examination of the psychometric characteristics of the math anxiety scales is critical at the item level. To our knowledge, however, previous research on the math anxiety scales has predominantly relied on factor analysis for validation (e.g., Baloğlu and Zelhart 2007; Beasley et al. 2001; Caviola et al. 2017; Erktin et al. 2006; Hopko 2003; Primi et al. 2014; Wigfield and Meece 1988), and such an approach is at the total score level based on the sum of the responses to all the items of a scale. Because this factor analysis approach does not provide information at the item level, it assumes equal contribution of different items to a scale, indicating an important disadvantage of this approach due to lack of determining the extent to which a single item contributes to the scale. Therefore, the purpose of this study was to examine the psychometric characteristics of the MANX, developed by Erol (1989), for measuring 8th grade students' math anxiety at the item level using the Rasch measurement model (specifically, the Rasch Rating Scale model). In particular, the present study is guided by the following research questions:

1- Does the internal structure of the MANX represent gradations of item difficulty?
2- Do the MANX items demonstrate acceptable model-data fit supporting the validity of inferences in terms of students' math anxiety levels?

The present study is one of the few studies (e.g., Prieto and Delgado 2007) that examine the validation of a math anxiety scale at the item level with the Rasch measurement model. Specifically, the present study is the first attempt to evaluate the item-level quality of the MANX with the Rasch Rating Scale model. We decided to use the Rasch Rating Scale model due to its advantages over other approaches such as factor analysis. First, if the data fit the Rasch model well, then person trait and item difficulty estimates can be obtained as independent from the particular sample (i.e., invariant measurement; Engelhard 2013). Second, the Rasch model transforms the ordinal raw data into interval measures (Wright and Mok 2004). However, approaches such as factor analysis rely on correlations of sample-dependent ordinal data, and it has the potential of contaminating the validity of a scale (Bond and Fox 2015). Third, the Rasch model provides a variable map that locates persons and items simultaneously on the latent trait so that one can distinguish the ordered structure of the items on the map. Hence, items and person trait levels can be compared directly on the variable map, with a great deal of overlap between the two indicating the degree of appropriateness of the sample for the scale. In addition to examining the psychometric characteristics of the MANX, this study provides a rationale for revising and improving its measurement qualities.

Past research has indicated that math anxiety appears to grow during the middle school years and to arrive at its peak during the early years of high school (Ashcraft and Moore 2009; Hembree 1990). Hill et al. (2016) considers the development of math anxiety as closely related to increasing educational demands when moving from lower grades to higher grades such as middle grades to secondary grades. Therefore, this study focuses on 8th grade students because this grade is a critical transition period from middle school to high school with potential high math anxiety levels, and because mathematics performance at this grade shapes the career path of Turkish students due to a nation-wide examination they need to take at the end of the year to gain acceptance to high schools in Turkey.

Method

Participants

The sample consists of 952 Turkish 8th grade students from nine schools in two cities located in different regions of Turkey. Students' range of age was between 13 and 15 years ($N = 478$ for females, $N = 474$ for males). Before the study was conducted, parental consent was obtained from one of the parents of every student.

Instruments

The Math Anxiety Scale (MANX), developed by Erol (1989), with 45 items was administered. The MANX has four response categories for each item including "never," "sometimes," "usually," and "always." The coefficient alpha (i.e., α), internal consistency reliability, was .91 in Erol (1989) and .92 in Erktin et al. (2006). By following the factors in Erktin et al., the coefficient alpha in the present study is .95 for the 45-item MANX ($\alpha = .91$ for *test and evaluation anxiety*, $\alpha = .87$ for *apprehension of math lessons*, $\alpha = .75$ for *use of mathematics in daily life*, and $\alpha = .68$ for *self-efficacy for mathematics*). While 36 items of the MANX are negatively worded, 9 items are positively worded. Before performing the analyses in this study, the 9 positively worded items were reverse-coded so that higher scores would demonstrate higher levels of math anxiety, and vice versa. Students completed the MANX during approximately 30 min of a class period. The English translation of the MANX and its underlying factors are presented in Appendix Table 2.

Data analysis

We evaluated the psychometric characteristics of the MANX based on the Rasch measurement model by assuming that students and items are located on a latent continuum of the trait being investigated (math anxiety in this study). The probability of agreeing with an item response category is determined by the location of a student on that continuum relative to the difficulty of that item (Rasch 1960/1980). Item difficulty represents students' propensity to agree with an item response category. For example, students who exhibit higher levels of math anxiety would be more likely to consider a response category of "never" as more difficult to endorse than "sometimes," a response category of "sometimes" more difficult to agree with than "usually," and so on. On the other hand, students with lower levels of math anxiety would be expected to consider "always" as more difficult to endorse than "usually," an item response category of "usually" more difficult to agree with than "sometimes," and so on. Because the data in this study were polytomous (having more than two response categories), and all the

MANX items have a fixed number of response categories (i.e., four response categories per item), we analyzed the data using the Rasch Rating Scale model (Andrich 1978) with the following equation:

$$P_{nix} = \frac{\exp\left[\sum_{j=0}^{x}\left(\theta_n - \left(\lambda_i + \delta_i\right)\right)\right]}{\sum_{x=0}^{m_i}\exp\left[\sum_{j=0}^{x}\left(\theta_n - \left(\lambda_i + \delta_i\right)\right)\right]} \tag{1}$$

where θ_n is the latent trait of a student, x is the item response category, λ_i is the item difficulty for a particular item, and δ_i is the category threshold parameter, which is assumed to be constant for all items. These estimates are reported in a logit (log odds) metric. Lower λ_i values (logits such as -1, -2, -3) indicate that students consider a particular item easier to endorse (i.e., higher levels of math anxiety), whereas higher λ_i values (logits such as 1, 2, 3) demonstrate that students have more difficulty agreeing with the item (i.e., lower levels of math anxiety).

To evaluate the psychometric characteristics of the MANX, the following analyses were performed. First, we evaluated the unidimensionality of the MANX by principal components analysis as implemented in the SPSS 16.0 software (SPSS Inc. 2007). Unidimensionality indicates that a single common trait explains the item responses. In reality, most of the scales are rarely strictly unidimensional because "unidimensionality is a relative matter—every human performance … is complex and involves a multitude of component abilities" (Andrich 1988, p. 9–10). In addition, dimensionality depends on the influence of both scale items and respondents who take the scale, and so may change from sample to sample. Therefore, it is important for researchers to focus on "essential unidimensionality" (Stout 1990), based on accepting scales as unidimensional if there exists one dominant trait observed in the data. In case of the scales being multidimensional, unidimensional models such as the Rasch measurement model have been found to do a good job in explaining multidimensional scale data in the presence of a dominant construct in the data (e.g., Anderson et al. 2017).

Second, we obtained Rasch person separation and item separation reliability values. A high reliability value (closer to 1) of Rasch person separation provides evidence for the sensitivity of the MANX in differentiating students with high versus low math anxiety. Similarly, a high reliability value (close to 1) of Rasch item separation demonstrates the distinctiveness of the items in their locations. Third, the variable map was evaluated in terms of the appropriateness of the sample for the MANX. Person traits being plotted on the same continuum as the item difficulties indicate the degree of appropriateness of the scale to the sample. While closeness between items and people demonstrate appropriateness, person traits being higher with respect to item difficulties indicate easiness of the scale, and vice versa. Fourth, we explored the model-data fit visually between the data and the model shown in Eq. (1). The alignment between the expected scores and the observed scores indicates good model-data fit. Fifth, we examined the item quality of the MANX based on the individual mean square (MS) fit statistics that show the extent to which each item represents the underlying trait (i.e., math anxiety). The ideal value of the MS fit statistics (Infit and Outfit) is 1.00 with standard deviations around 0.20 (Engelhard 2009). While an MS fit statistic (Infit and Outfit) greater than 1.00 indicates more variation in students' responses to that particular item (i.e., unexpected responses), an MS fit statistic smaller than 1.00 provides evidence that the item has less variation in students' responses (Linacre 2005). In this study, an MS fit statistic between 0.60 and 1.40 was considered a reasonable fit for the MANX items (Linacre and Wright 1994). All analyses except the analysis of unidimensionality were performed using the Facets program (Linacre 2013), which provides estimates for each item and student in addition to summary statistics for all the items and students.

Results

Assessment of unidimensionality for the MANX

The unidimensionality of the MANX was confirmed with a principal component analysis (Kaiser-Meyer-Olkin measure of sampling adequacy = .961, Barlett's Test of Sphericity, $p < .01$) reporting eigenvalues of the first three components as 14.3, 2.6, and 1.9. The total variance explained by the first component was 31.7%. Because the amount of variance explained by the first component exceeds 20% (Reckase 1979), the MANX was concluded to be essentially unidimensional. Moreover, the variance explained by the Rasch dimension using the Facets program was 35.7, indicating another evidence that the MANX is essentially unidimensional in measuring a single, unified trait (i.e., math anxiety).

Rasch person and item separation reliability

Table 1 presents Rasch person separation and item separation reliability values. Based on Table 1, the MANX has a high Rasch person separation reliability with Rel = .93, χ^2 (951) = 11,066.3, and $p < .01$, indicating that the scale was able to discriminate between students with varying levels of math anxiety. The MANX also has a high Rasch item separation reliability with Rel = .99, χ^2 (44) = 6462.4, $p < .01$, implying that the items have a good spread in terms of item ordering and hierarchy.

Variable map

Figure 1 displays a variable map that shows the item locations of the MANX items and the person locations for each student who completed the survey. These locations are estimated in logits (log odds units) as the unit of measures. This means that the higher the logit value for an item is, the less likely it was for a person to endorse that particular item. On the other hand, higher logit values for each person indicate higher levels of math anxiety. On the variable map, the mean item location is constrained to be 0.00 with person locations being relative to that mean item location. While column 1 shows the latent continuum in terms of logit values as the unit of measurement underlying the MANX, columns 2 and 3 represent the locations of the students and the items. Moreover, in column 3, the item locations are displayed with four sub-columns that show the underlying factors ranging from *test and evaluation anxiety* (i.e., sub-column 1) to *self-efficacy for mathematics* (i.e., sub-column 4) as left to right.

Table 1 Summary statistics for students and items

	Students	Items
Measures		
M	− 0.58	0.00
SD	0.79	0.47
N	952	45
Infit		
M	1.03	1.03
SD	0.41	0.23
Outfit		
M	1.02	1.02
SD	0.45	0.32
Rasch person/item separation reliability	0.93	0.99
χ^2 statistic	11,066.3*	6462.4*
Degrees of freedom	951	44

*$p < .01$

According to the variable map, there exists a good overlap between the person trait and item difficulty due to the match between the mean of the person trait ($M = -0.58$) and the mean of the item difficulty ($M = 0.00$). This suggests that the majority of the items were appropriate for the sample. While person trait measures range from 2.13 logits to − 5.41 logits ($M = -0.58$, SD = 0.79, $N = 952$), the item difficulties range from 1.14 logits to − 1.08 logits ($M = 0.00$, SD = 0.47, $N = 45$). The most difficult item for students to endorse is item 9, "I feel confused when I try to count the change I received from my purchase from the school cafeteria; most of the time I just get what is given to me without counting the change," with the location at 1.14 logits (also reported in Appendix Table 2). The easiest item to agree with, however, is item 10, "I would like to keep the accounts for a school club or activity that I am participating in", with − 1.08 logits.

Regarding the variable map, the MANX appears to have a few limitations. First, the MANX has a lack of items that provide information about students who are located at the very high end of the continuum, above 1.14 logits, and those who are placed on the very low end of the continuum, particularly below − 0.85 logits. Hence, the MANX seems to be not targeted at identifying students who have very high and low levels of math anxiety. Second, in terms of the distributions of items with respect to their underlying factors (see Fig. 1 and Appendix Table 2), most of the items are spread sufficiently between 1.14 logits and − 1.08 logits except the items which belong to *use of mathematics in daily life*, represented by seven items (i.e., items 9, 15, 17, 26, 29, 38, 45). That five of these seven items (i.e., items 9, 15, 17, 26, 29) are located at the high end of the continuum indicates that these items are the most difficult items for students to agree with. This, in turn, suggests that these items do not provide information about students with low or moderate levels of math anxiety related to the *use of mathematics in daily life* on the high end of the continuum. Third, in terms of the redundancy of the items that provide the exact same information, item 17 and item 26, represented by the *use of mathematics in daily life*; item 6 and item 37, represented by *apprehension of math lessons*; and item 21 and item 44,

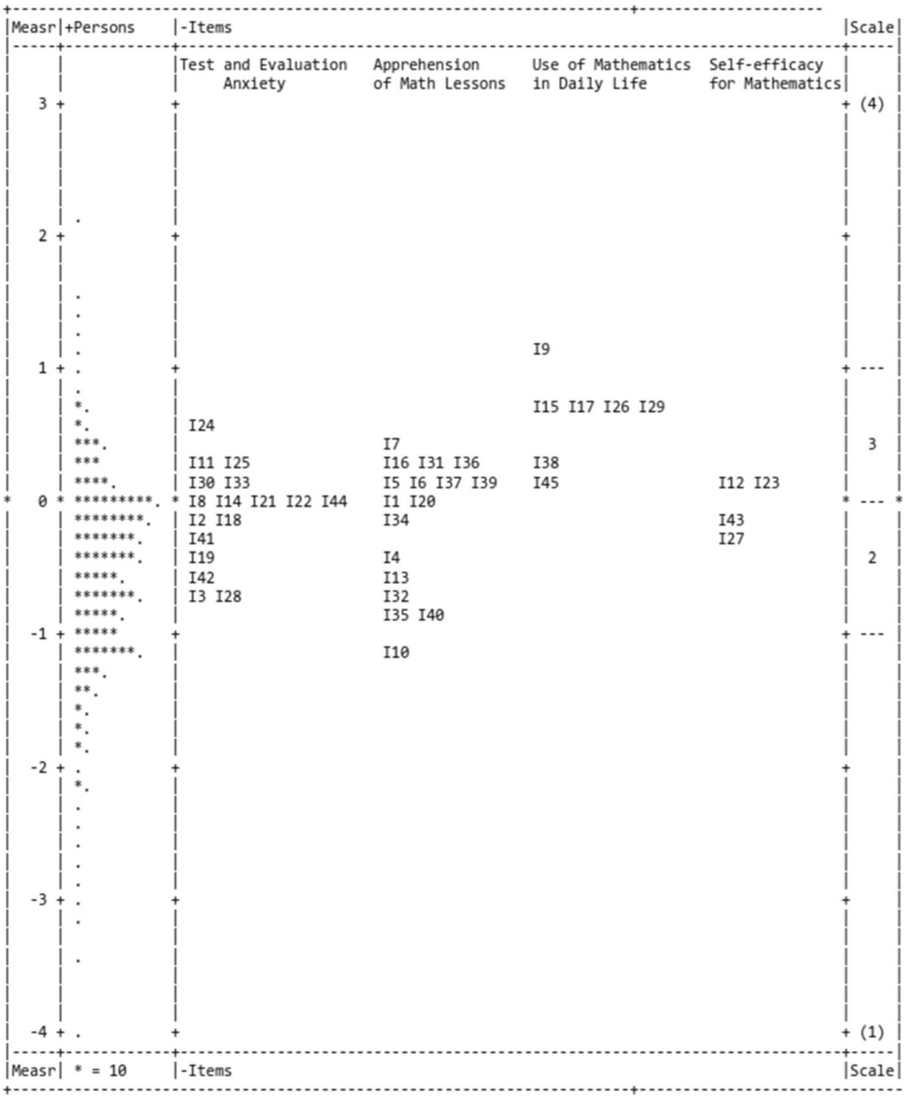

Fig. 1 Variable map of students' math anxiety levels and item locations

represented by *test and evaluation anxiety*, have the same mean ratings, item difficulties, and standard errors as within pairs (see Appendix Table 2). Therefore, item 17, item 37, and item 21, which have slightly worse fit statistic (Infit and Outfit), can be eliminated from the scale to shorten the length of the scale's administration.

Model-data fit

Figure 2 presents the graphical display of the model-data fit between the data and the model shown in Eq. (1). As opposed to item response theory (IRT) models such as the two-parameter models and the three-parameter models that attempt to fit the particular model to the data, the Rasch measurement model necessitates the fit of the data to the model with predetermined requirements (Andrich 2004). Specifically, the Rasch measurement model evaluates whether the data fit the model by comparing the response pattern observed in the data to the expected pattern by the model (Bond and Fox 2015).

Based on Fig. 2, the model-data fit is reasonable through close alignment between the expected and observed values despite outliers on the high end of the continuum. Hence, the data of this study appear to fit the Rasch Rating Scale model, but the MANX can also be strengthened through some modifications of the items for coping with the outliers in the figure. Furthermore, the figure provides a strong evidence for the appropriateness of using a variable map to represent gradations of item difficulties and person traits simultaneously on the continuum of math anxiety. In addition, the average logit measures for each category ranging from "never" to "always" are −0.36, −0.10, 0.06, and 0.40, respectively. This provides another evidence for good model-data fit due to the increase in average logit measures by response category.

Item quality

According to the summary statistics shown in Table 1, even though both students and items have slightly more variation than expected for the Outfit statistic ($M = 1.02$, SD = 0.45 for students; $M = 1.02$, SD = 0.32 for items), these observed values for the MS fit

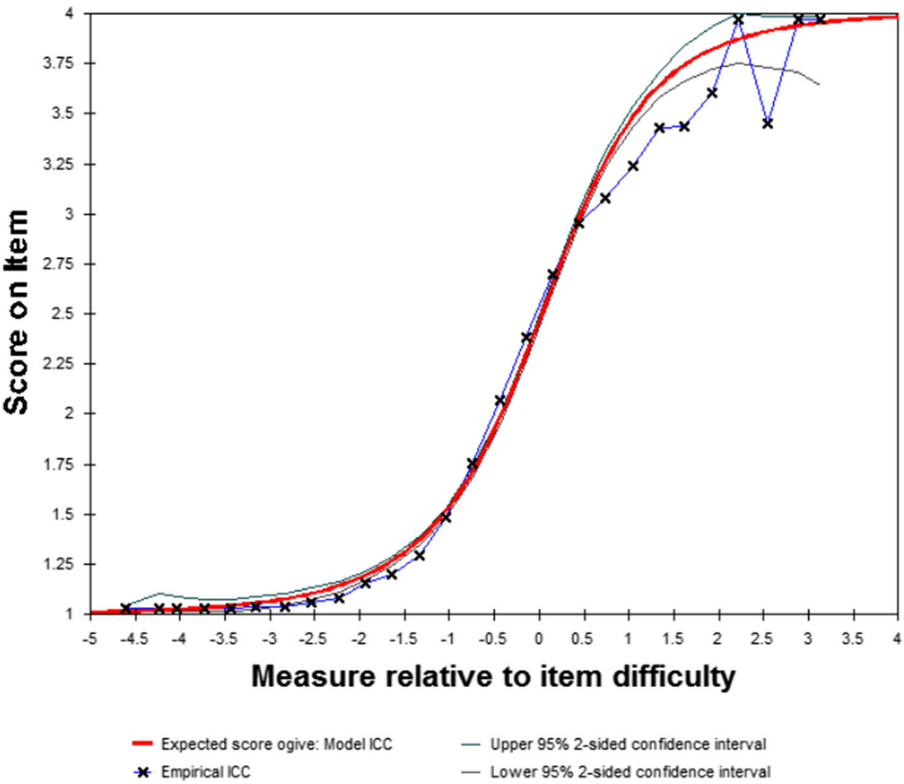

Fig. 2 Model-data fit for the Rasch Rating Scale model. ICC = item characteristic curve

statistic (Infit and Outfit) can be considered as close to the expected values (i.e., the MS fit statistic is 1.00 with standard deviations around 0.20 (Engelhard 2009).

Appendix Table 2 presents item quality indices for each item, and these indices indicate that most of the items (i.e., 37 of the 45 items) have a good fit regarding the range of the MS fit statistic between 0.60 and 1.40 (Linacre and Wright 1994). Eight of the 45 items, however, demonstrate misfit including Item 5, Item 9, Item 10, Item 13, Item 20, Item 34, Item 35, and Item 43. Item 34, “I feel confused in math class,” is the only item in the scale that exhibits less variation than expected (Infit = 0.59 and Outfit = 0.58), indicating that item 34 does not provide information to distinguish between students with varying levels of math anxiety due to the inconsistent response patterns underlying this item. That is, two students with different math anxiety levels can respond to item 34 in the same way.

The remaining seven misfit items have more variation than expected due to the unpredictability of students’ responses in these items. Five of these items (i.e., items 10, 13, 20, 35, 43) are positively worded (reverse-coded). Item 10, “I would like to keep the accounts for a school club or activity that I am participating in,” is the easiest item to agree with for students. Misfit in this item appears to be related to some high anxious students’ unexpected low scores on this item. This misfit might happen because this item seems to measure the degree that one is interested in taking responsibility instead of the degree of apprehension of math lessons. Similarly, item 20, “I always enter the first math class of the year with hope,” is another item that shows misfit. This item also seems to target a different dimension than apprehension of math lessons, which is based on the degree of one’s hope or expectation that he or she will experience about mathematics throughout the year. These two items (i.e., item 10, 20) were also found to be the least correlated items to the total raw score of the MANX in the study by Erktin et al. (2006). Next, item 13, “I like any topic explained to me in numbers and graphs rather than verbal explanations,” demonstrates misfit. This misfit based on unexpected variation in students’ responses might be due to the fact that this item is positively worded (reverse-coded). In other words, item 13 and its negation might not have the same meaning or interpretation for the students in this study, leading to response style bias. The same explanation also holds for the presence of unexpected variation in students’ responses to item 35 and item 43. Because both Infit and Outfit statistics of these five positively worded items are outside of acceptable range, they can be considered as fundamentally flawed (Ludlow et al. 2014) and need to be removed from the scale. Although placing both negatively and positively worded items into a scale increases random response variation, it has a high potential to contaminate the validity of the items.

Furthermore, one of the remaining two misfit items, which are not reverse-coded, is item 9, targeting the *use of mathematics in daily life*. This misfit indicates that even if a student rated high on this item, he or she would not necessarily endorse easier items related to the *use of mathematics in daily life* such as item 38, “I feel bothered by the necessity of making calculations by solving mathematical problems in my daily life even if they are simple,” and item 45, “If a friend asks me to solve a math problem that has been published in a magazine, I am afraid of being embarrassed by not being able to solve even the easiest problems.” (see Fig. 1). Given that item 9 is the most difficult item to endorse for students, the reason for this misfit seems to be related to some low anxious students’ unexpected high scores on this item. The last misfit item is item 5, “I dislike the formulas in the science classes.” In this item, more variation in the student responses than expected might stem from the fact that students do not necessarily perceive mathematics when they see the word “formulas” in this item.

Discussion

The purpose of this study was to assess the psychometric characteristics of the MANX with a sample of 8th grade Turkish students at the item level using the Rasch Rating Scale model. Although most of the studies in Turkey have used the MANX to measure math anxiety levels of middle school and high school students (e.g., Ader and Erktin 2010; Bekdemir 2010; Erden and Akgül 2010), very limited research about the psychometric characteristics of the MANX exists, including its underlying factor structure (e.g., Erktin et al. 2006). Considering the fact that factor analysis results are sample dependent and might cause misleading conclusions about the validity of the scales (Bond and Fox 2015), the present study contributes to our understanding about the psychometric characteristics of the MANX using the Rasch Rating Scale model.

Given that the process of validating a scale is usually seen as “the most unsatisfactory aspect” (Schilling and Hill 2007, p. 70) of scale development, validated scales are necessary for conducting high-quality research and for applicability in wider contexts. It is important to emphasize that a scale being widely used does not necessarily guarantee that the scale is reliable and measures the construct that is intended to measure. A scale which has not been validated yet has the great risk of not providing accurate information about respondents’ trait levels, and this in turn might cause misleading decisions on providing particular early intervention efforts to those respondents. Therefore, scales need to be validated in order to make sure that they provide accurate information about their intended use(s) and to apply them in wider contexts.

In this study, we specifically examined unidimensionality, Rasch person/item separation reliability, model-data fit, distribution of the variable map based on the locations of persons and items, and individual item qualities of the MANX. In terms of dimensionality, the MANX was found to be essentially unidimensional. In terms of Rasch person/item separation reliability, the MANX was found to be highly reliable for persons and items, indicating that most of the items are sensitive enough to capture different levels of math anxiety, and most students’ observed values overlap with the expected values. Similarly, there was a close correspondence between the data and the Rasch Rating Scale model except for the outliers in the high ends of the continuum, indicating a reasonable model-data fit.

Regarding the variable map, results from the present study indicate that the MANX, overall, captures a range of item difficulties that reflect the order structure of math anxiety. However, we also found that items related to the *use of mathematics in daily life* are quite difficult for students to agree with, and there is a shortage of items at the very high and low ends of the continuum across the four factors. Therefore, it can be concluded that the MANX is able to capture students with moderate levels of math anxiety, but it is not sensitive enough to discriminate among students with very high and low levels of math anxiety. For example, based on the Rasch Rating scale model analysis in the present study, a student with a measure of −3.00 logits and another student with a measure of −2.00 logits would be more likely to respond to the MANX in the same way, because there is no item on the scale that targets very low levels of math anxiety. Hence, we would not be able to detect which student has relatively more or less math anxiety in comparison to each other. Similarly, it would not be possible to distinguish math anxiety levels between a student with a measure of 3.00 logits and another student with a measure of 2.00 logits, and decide who possesses more math anxiety levels due to lack of item at the high end of the continuum.

Finally, in terms of item quality, 37 of the 45 items exhibited good psychometric quality, whereas eight items were misfits. Among the eight misfit items, five of them were reverse-coded. It is quite possible that these reverse-coded items confused the students and contributed to bias toward response styles, influencing the validity of the MANX. In addition, three items were found to be redundant because they provided the exact same information with identical pairs of each item, displaying the same contributions to the sensitivity of the MANX. Hence, they can be removed from the scale to reduce the length of the administration.

Use of the Rasch measurement model in the present study provided a more powerful analysis of the functioning of the MANX at the item level and offered in-depth insights on applicability of the scale in wider contexts due to its sample independent nature. Whereas factor analytic approaches such as exploratory factor analysis and confirmatory factor analysis do not provide invariant measurement due to their sample dependency, the Rasch measurement model is independent from the items that form the scale and independent from the sample that the scale is administered to. This feature will increase the applicability of the scale in wider contexts and replicability of the present study in different populations due to the possibility of direct application of the standards developed in one study to data administered in another study (Yan and Mok 2012). The present study contributes to an awareness and effectiveness of the use of the Rasch measurement model in validating scales for measuring affects such as math anxiety.

In conclusion, the results of this study provide strong evidence for the validation of the MANX despite the need for considering the deletion of eight items that are misfits and of three items that are redundant. The revision or development of new items can also be considered to capture gradations of challenges at the very high and low ends of the continuum to detect students with very high and low levels of math anxiety.

Implications and future directions

This study presents the validation of the MANX using the Rasch measurement model, for assessing math anxiety levels of 8th grade students. The findings have important implications for teachers, counselors, and educators who are interested in increasing students' success in mathematics. Given that success in mathematics is highly negatively associated with math anxiety (Hembree 1990), validated math anxiety scales with good psychometric characteristics such as the MANX can provide a tool for teachers, counselors, and educators to diagnose students with different levels of math anxiety accurately and target those students with early intervention efforts. Moreover, the finding that the majority of the MANX items were of high quality implies that the items provide adequate information to measure moderate levels of students' math anxiety. However, the MANX was found not to be able to distinguish students with very high and low levels of math anxiety. Hence, teachers, counselors, and educators who are using the MANX in its current form should be cautious when interpreting their results for students who have very high and low math anxiety levels.

In addition to the MANX scale being a self-report scale, limitations associated with this study include the grade level. Because this study consists of responses of students from one grade level (i.e., 8th grade), the results might not be generalizable to students in other grades. There are a few suggestions for future studies. First, future research should consider eliminating the 11 items that either demonstrated misfit or were redundant, and include only the remaining 34 items that have good item qualities. Second, there is a need for the revision or development of new items to capture students with very high and low levels of math anxiety. Third, permanent changes to the MANX need to be made after conducting pilot studies with other 8th grade samples. Fourth, in terms of generalizability of the findings, this study should be replicated by future studies with the Rasch measurement model analysis on samples of 6th and 7th grade students, and of high school students. Last, future research should prioritize using the Rasch model for scale development and validation studies over other approaches, such as factor analysis, due to its advantages for revealing the characteristics of the items as independent from the sample.

Funding The authors received no financial support for the research, authorship, and/or publication of this article.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Appendix

Table 2 Item information and item quality index in the Rasch analysis of the MANX

Items	Content	Underlying dimension	Mean rating	Measure	SE	Infit MS	Outfit MS
Item 1	When my friend is asked to answer a question in Math class, I feel glad that I am not in his/her shoes.	Apprehension of Math Lessons	1.81	0.06	0.04	0.72	0.80
Item 2	I panic when I start the Math part of a common test.	Test and Evaluation Anxiety	2.05	−0.20	0.03	0.78	0.76
Item 3	I am intimidated when asked to answer a question that I do not completely know the answer to.	Test and Evaluation Anxiety	2.56	−0.65	0.03	0.79	0.82
Item 4	I like doing Math homework.***	Apprehension of Math Lessons	2.38	−0.49	0.03	0.89	0.91
Item 5	I dislike the formulas in the science classes.		1.73	0.15	0.04	1.31	*1.67*

Table 2 (continued)

Items	Content	Underlying dimension	Mean rating	Measure	SE	Infit MS	Outfit MS
		Apprehension of Math Lessons					
Item 6	I panic when I am assigned homework which includes a lot of math problems.	Apprehension of Math Lessons	1.77	0.10	0.04	0.86	0.80
Item 7	I feel butterflies in my stomach when I prepare to study a hard math topic.	Apprehension of Math Lessons	1.52	0.46	0.04	1.03	0.89
Item 8	I become unable to think about anything one hour prior to math exam.	Test and Evaluation Anxiety	1.84	0.02	0.04	1.01	1.04
Item 9	I feel confused when I try to count the change I received from my purchase from the school cafeteria; most of the time I just get what is given to me without counting the change.	Use of Mathematics In Daily Life	1.25	1.14	0.06	*1.42*	1.23
Item 10	I would like to keep the accounts for a school club or activity that I am participating in.***	Apprehension of Math Lessons	3.04	− 1.08	0.04	*1.41*	*1.80*
Item 11	I feel intimidated to check my math score when I am given my class report.	Test and Evaluation Anxiety	1.60	0.33	0.04	0.99	0.81
Item 12	I feel reluctant to explain the problems even the Self-efficacy for ones I can solve.	Self-efficacy for Mathematics	1.73	0.16	0.04	0.91	0.84
Item 13	I like any topic explained to me in numbers and graphs rather than verbal explanations.***	Apprehension of Math Lessons	2.51	− 0.60	0.03	*1.43*	*1.77*
Item 14	I feel terrible a day before the math exam.	Test and Evaluation Anxiety	1.86	0.00	0.04	0.81	0.74
Item 15	Even if I think that a shop clerk gave me the wrong change, I say nothing because I cannot make calculations while somebody is watching me.	Use of Mathematics In Daily Life	1.38	0.75	0.05	1.30	1.24
Item 16	The math book bothers me.	Apprehension of Math Lessons	1.58	0.36	0.04	0.98	0.87
Item 17	I cannot even make an addition operation while somebody is watching.	Use of Mathematics In Daily Life	1.42	0.65	0.04	1.28	1.10
Item 18	I become so nervous before the important math exams that I forget all I know.	Test and Evaluation Anxiety	2.01	− 0.15	0.03	0.83	0.80
Item 19	I feel afraid when the teacher gives a pop quiz on math.		2.26	− 0.39	0.03	0.87	0.81

Table 2 (continued)

Items	Content	Underlying dimension	Mean rating	Measure	SE	Infit MS	Outfit MS
		Test and Evaluation Anxiety					
Item 20	I always enter the first math class of the year with hope.***	Apprehension of Math Lessons	1.80	0.07	0.04	*1.62*	*1.71*
Item 21	While studying for a math exam, I may not prepare enough at times because of worrying about the score I will get.	Test and Evaluation Anxiety	1.81	0.06	0.04	0.78	0.75
Item 22	I feel an inability to succeed while going through the pages of the math book.	Test and Evaluation Anxiety	1.88	−0.02	0.04	0.75	0.70
Item 23	I cannot dare to ask the points I do not get in the math class.	Self-efficacy for Mathematics	1.77	0.10	0.04	0.92	0.84
Item 24	I feel uneasy even when I calculate the GPA for my class report.	Test and Evaluation Anxiety	1.48	0.53	0.04	0.92	0.78
Item 25	I feel uneasy a week before the math exam.	Test and Evaluation Anxiety	1.62	0.30	0.04	0.91	0.77
Item 26	Even making calculation related to time gives me discomfort.	Use of Mathematics In Daily Life	1.42	0.65	0.04	1.09	1.04
Item 27	I can easily ask something that I did not understand to the math teacher after the class.***	Self-efficacy for Mathematics	2.14	−0.28	0.03	1.16	1.11
Item 28	I feel nervous and pessimistic while waiting for the announcement of the result for a math exam that I think I failed at.	Test and Evaluation Anxiety	2.63	−0.71	0.03	1.13	1.25
Item 29	When I am asked to help a primary school student with his/her homework, I may refuse to help because I feel afraid that there may be some problems I couldn't solve.	Use of Mathematics In Daily Life	1.41	0.67	0.04	1.21	1.05
Item 30	When I think of the math subjects I have to learn before graduating from high school, I doubt if I am ever going to finish school.	Test and Evaluation Anxiety	1.71	0.18	0.04	0.93	0.87
Item 31	Dealing with numbers makes me annoyed.	Apprehension of Math Lessons	1.66	0.24	0.04	0.91	0.89
Item 32	Geometry questions remind me of fun puzzles.***	Apprehension of Math Lessons	2.69	−0.75	0.03	1.25	1.38
Item 33	I feel tense when my friend solves a problem and I cannot understand his/her solution.		1.73	0.16	0.04	1.09	1.13

Table 2 (continued)

Items	Content	Underlying dimension	Mean rating	Measure	SE	Infit MS	Outfit MS
		Test and Evaluation Anxiety					
Item 34	I feel confused in math class.	Apprehension of Math Lessons	2.05	−0.20	0.03	*0.59*	*0.58*
Item 35	The most likable part of the social classes are the parts that consist of math, even if they are miniscule.***	Apprehension of Math Lessons	2.80	−0.85	0.04	1.39	*1.74*
Item 36	I struggle with listening to the teacher in the math class.	Apprehension of Math Lessons	1.62	0.30	0.04	0.86	0.81
Item 37	I feel uneasy when I know that the following lesson is math.	Apprehension of Math Lessons	1.77	0.10	0.04	0.87	0.78
Item 38	I feel bothered by the necessity of making calculations by solving mathematical problems in my daily life even if they are simple.	Use of Mathematics In Daily Life	1.65	0.27	0.04	0.85	0.74
Item 39	I feel depressed by the math book.	Apprehension of Math Lessons	1.68	0.21	0.04	0.87	0.75
Item 40	Opening any book on math and looking at one of its pages full of mathematical problems makes me happy.***	Apprehension of Math Lessons	2.78	−0.83	0.04	1.10	1.20
Item 41	When I am given a problem to solve, I panic if I cannot remember the formula necessary for the solution.	Test and Evaluation Anxiety	2.21	−0.35	0.03	0.95	1.01
Item 42	Five minutes before the math exam, my heart starts starts beating fast.	Test and Evaluation Anxiety	2.39	−0.50	0.03	1.16	1.23
Item 43	When I think that I succeeded at a math exam, I feel relaxed and peaceful while waiting for the announcement of the results.***	Self-efficacy for Mathematics	1.99	−0.14	0.03	1.37	*1.45*
Item 44	If the teacher asks me to solve a math problem that I have been working on for a while at the blackboard, I forget what I have done out of excitement.	Test and Evaluation Anxiety	1.81	0.06	0.04	0.93	0.86
Item 45	If a friend asks me to solve a math problem that has been published in a magazine, I am afraid of being embarrassed by not being able to solve even the easiest problems.	Use of Mathematics In Daily Life	1.73	0.15	0.04	1.01	0.88

***Reverse-coded. *MANX* Math Anxiety Scale (Erol 1989); *MS* mean square. Italics indicates MS fit statistic out of the range between 0.60 and 1.40

References

Ader, E., & Erktin, E. (2010). Coping as self-regulation of anxiety: a model for math achievement in high-stakes tests. *Cognition, Brain, Behavior: An Interdisciplinary Journal, 14*(4), 311–332.

Alexander, L., & Martray, C. (1989). The development of an abbreviated version of the Mathematics Anxiety Rating Scale. *Measurement and Evaluation in Counseling and Development, 22*, 143–150.

Anderson, D., Kahn, J. D., & Tindal, G. (2017). Exploring the robustness of a unidimensional item response theory model with empirically multidimensional data. *Applied Measurement in Education, 30*(3), 163–177.

Andrich, D. (1978). Rating formulation for ordered response categories. *Psychometrika, 43*, 561–573.

Andrich, D. (1988). *Rasch models for measurement*. Newbury Park, CA: Sage.

Andrich, D. (2004). Controversy and the Rasch model: a characteristic of incompatible paradigms? *Medical Care, 42*, 1–16.

Ashcraft, M. H. (2002). Math anxiety: personal, educational, and cognitive consequences. *Current Directions in Psychological Science, 11*(2), 181–185.

Ashcraft, M. H., & Moore, A. M. (2009). Mathematics anxiety and the affective drop in performance. *Journal of Psychoeducational Assessment, 27*(3), 197–205.

Aydin, E., Delice, A., Dilmaç, B., & Ertekin, E. (2009). The influence of gender, grade and institution on primary school mathematics student teachers' anxiety levels. *Elementary Education Online, 8*(1), 231–242.

Baloğlu, M., & Zelhart, P. F. (2007). Psychometric properties of the revised mathematics anxiety rating scale. *The Psychological Record, 57*, 593–611.

Beasley, T. M., Long, J. D., & Natali, M. (2001). A confirmatory factor analysis of the mathematics anxiety scale for children. *Measurement and Evaluation in Counseling and Development, 34*, 14–26.

Bekdemir, M. (2010). The pre-service teachers' mathematics anxiety related to depth of negative experiences in mathematics classroom while they were students. *Educational Studies in Mathematics, 75*, 311–328.

Bindak, R. (2005). İlköğretim öğrencileri için matematik kaygı ölçeği (Math anxiety scale for elementary school students). *Fırat Üniversitesi, Fen ve Mühendislik Bilimleri Dergisi, 17*(2), 442–448.

Bond, T. G., & Fox, C. M. (2015). *Applying the Rasch model: fundamental measurement in the human sciences*. New York, NY: Routledge.

Caviola, S., Primi, C., Chiesi, F., & Mammarella, I. C. (2017). Psychometric properties of the Abbreviated Math Anxiety Scale (AMAS) in Italian primary school children. *Learning and Individual Differences, 55*, 174–182.

Engelhard, G. (2009). Using item response theory and model-data fit to conceptualize differential item and person functioning for students with disabilities. *Educational and Psychological Measurement, 69*(4), 585–602.

Engelhard, G. (2013). *Invariant measurement: using Rasch models in the social, behavioral, and health sciences*. New York, NY: Routledge.

Erden, M., & Akgül, S. (2010). Predictive power of math anxiety and perceived social support from teacher for primary students' mathematics achievement. *Journal of Theory and Practice in Education, 6*(1), 3–16.

Erktin, E., Dönmez, G., & Özel, S. (2006). Psychometric characteristics of the mathematics anxiety scale. *Education and Science, 31*(140), 26–33.

Erol, E. (1989). *Prevalence and correlates of math* anxiety *in Turkish high school students*. Unpublished master thesis, Bogazici University.

Fennema, E., & Sherman, J. A. (1976). Fennema-Sherman Mathematics Attitude Scale: instruments designed to measure attitudes toward the learning of mathematics by males and females. *JAS Catalog of Selected Document of Psychology, 6*, 31.

Hembree, R. (1990). The nature, effects, and relief of mathematics anxiety. *Journal for Research in Mathematics Education, 21*, 33–46.

Hill, F., Mammarella, I. C., Devine, A., Caviola, S., Passolunghi, M. C., & Szücs, D. (2016). Math anxiety in primary and secondary school students: gender differences, developmental changes and anxiety specificity. *Learning and Individual Differences, 48*, 45–53.

Ho, H., Senturk, D., Lam, A. G., Zimmer, J. M., Hong, S., Okamoto, Y., & Chiu, S. (2000). The affective and cognitive dimensions of math anxiety: a cross-national study. *Journal for Research in Mathematics Education, 31*(3), 362–379.

Hopko, D. R. (2003). Confirmatory factor analysis of the Math Anxiety Rating Scale-Revised. *Educational and Psychological Measurement, 63*(2), 336–351.

Linacre, J. M. (2005). *A user's guide to WINSTEPS*. Chicago: Winsteps.com.

Linacre, J. M. (2013). Facets Rasch Measurement (Version 3.71.3) [Computer Software]. Chicago, IL: Winsteps.com.
Linacre, J. M., & Wright, B. D. (1994). Reasonable mean-square fit values. *Rasch Measurement Transactions, 8*(3), 370.
Ludlow, L. H., Matz-Costa, C., Johnson, C., Brown, M., Besen, E., & James, J. B. (2014). Measuring engagement in later life activities: Rasch-based scenario scales for work, caregiving, informal helping, and volunteering. *Measurement and Evaluation in Counseling and Development, 47*(2), 127–149.
Ma, X. (1999). Meta-analysis of the relationship between anxiety toward mathematics and achievement in mathematics. *Journal for Research in Mathematics Education, 30*, 520–540.
Prieto, G., & Delgado, A. R. (2007). Measuring math anxiety (in Spanish) with the Rasch Rating Scale Model. *Journal of Applied Measurement, 8*, 149–160.
Primi, C., Busdraghi, C., Tomasetto, C., Morsanyi, K., & Chiesi, F. (2014). Measuring math anxiety in Italian college and high school students: Validity, reliability and gender invariance of the Abbreviated Math Anxiety Scale (AMAS). *Learning and Individual Differences, 34*, 51–56.
Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen, Denmark: The Danish Institute of Education Research. Expanded edition (1980) with foreword and afterword by B.D. Wright. Chicago, IL: The University of Chicago Press).
Reckase, M. D. (1979). Unifactor latent trait models applied to multifactor tests: results and implications. *Journal of Educational Statistics, 4*, 207–230.
Richardson, F. C., & Suinn, R. M. (1972). The mathematics anxiety rating scale: psychometric data. *Journal of Counseling Psychology, 19*(6), 551–554.
Schilling, S. G., & Hill, H. C. (2007). Assessing measures of mathematical knowledge for teaching: a validity argument approach. *Measurement, 5*(2–3), 70–80.
SPSS Inc. (2007). SPSS for Windows, Version 16.0. Chicago, SPSS Inc.
Stout, W. F. (1990). A new item response theory modeling approach with applications to unidimensional assessment and ability estimation. *Psychometrika, 55*, 293–326.
Suinn, R. M., & Edwards, R. (1982). The measurement of mathematics anxiety. The Mathematics Anxiety Rating Scale for Adolescents–MARS-A. *Journal of Clinical Psychology, 38*(3), 576–580.
Suinn, R., Taylor, S., & Edwards, R. (1988). Suinn Mathematics Anxiety Rating Scale for Elementary School Students (MARS-E). Psychometric and normative data. *Educational and Psychological Measurement, 48*, 979–986.
Wigfield, A., & Meece, J. L. (1988). Math anxiety in elementary and secondary school students. *Journal of Educational Psychology, 80*, 210–216.
Wright, B. D., & Mok, M. M. C. (2004). An overview of the family of Rasch measurement models. In E. Smith & R. Smith (Eds.), *Introduction to Rasch measurement* (pp. 1–24). Maple Grove, MN: JAM Press.
Yan, Z., & Mok, M. M. C. (2012). Validating the coping scale for Chinese athletes using multidimensional Rasch analysis. *Psychology of Sport and Exercise, 13*, 271–279.
Young, C. B., Wu, S. S., & Menon, V. (2012). The neurodevelopmental basis of math anxiety. *Psychological Science, 23*(5), 492–501.