Revista de Psicodidáctica, 2014, 19(1), 67-91
www.ehu.es/revista-psicodidactica
ISSN: 1136-1034 eISSN: 2254-4372

DOI: 10.1387/RevPsicodidact.8961

Attitudes Towards Mathematics: Construction and Validation of a Measurement Instrument

Andrés Palacios*, Víctor Arias**, and Benito Arias*

University of Valladolid (Spain)*, University of Talca (Chile)**

Abstract

The measure of attitudes towards mathematics is a valuable area within the so-called affective domain in mathematics due to the number of investigations and the extension of their relations. However, most of the instruments currently available to measure these attitudes are validated by not overly robust psychometric procedures and, sometimes, in not very large samples. Using a sample of 4,807 students of all the non-university levels and following both the classical test theory, structural equation models (measurement models), and the proposal of the item response theory (graded response model), a solid and robust instrument to measure attitudes towards mathematics is presented, with contrasted evidence of validity and reliability.

Keywords: Attitudes towards mathematics, exploratory factor analysis, confirmatory factor analysis, graded response model, psychometrics.

Resumen

La medida de las actitudes hacia las matemáticas supone un campo de gran valor dentro de lo que se conoce como dominio afectivo matemático por el número de investigaciones y por la amplitud de sus relaciones. No obstante, los instrumentos disponibles en la actualidad para medir estas actitudes están en la mayoría de los casos validados mediante procedimientos psicométricos poco robustos y, en algunas ocasiones, con tamaños muestrales no muy elevados. A partir de una muestra de 4.807 alumnos de todos los niveles no universitarios y siguiendo tanto el acercamiento de la Teoría Clásica de los Test como los modelos de ecuaciones estructurales (modelos de medida) y el planteamiento de la Teoría de Respuesta a los Ítems (modelo de respuesta graduada) se presenta un instrumento de medida de las actitudes hacia las matemáticas sólido y robusto y con evidencias contrastadas de validez y fiabilidad.

Palabras clave: Actitudes hacia las matemáticas, análisis factorial exploratorio, análisis factorial confirmatorio, modelo de respuesta graduada, psicometría.

Acknowledgments: This research was also supported by the Spanish Ministry of Science and Innovation (EDU2009-12063).

Correspondence concerning this article should addressed to Andrés Palacios Picos. Escuela Universitaria de Magisterio. Plaza Alto de los Leones de Castilla. Campus María Zambrano. 40005 Segovia. E-mail: palacios@psi.uva.es

Introduction

The works of McLeod (1988, 1992) on affect in mathematics are an inflection point in the research in which, till this time, rational and cognitive aspects were predominant. In one of his works (McLeod, 1988) which, in a sense, marks the beginning of concern with emotions and feelings in mathematics, he established a distinction —now classical— between attitudes, beliefs, and emotions as the components of what is now known as the *affective domain in mathematics*. Among these components, attitudes have played a predominant role in mathematical education due to the number of investigations they have generated.

Gil, Blanco, and Guerrero (2005) note that, in the world of mathematics, this concept of attitude has been employed with a definition that is not as clear as the one used in Psychology, as a predisposition with an emotional charge that directs and/or influences behavior; a definition that underlines three basic components of attitudes: cognition or beliefs about the target of the attitude, affect or the evaluative charge of such beliefs, and a behavioral intention toward the attitude.

Nevertheless, with regard to mathematics, we can distinguish *mathematical attitudes* and *attitudes towards mathematics*. *Attitudes towards mathematics* refer to the valuation, the appraisal, and the enjoyment of this discipline, underlining the affective facet more than the cognitive one. *Mathematical attitudes*, in contrast, refer to the way one uses general capacities that are relevant for mathematics (such as mental openness, flexibility when seeking solutions to a problem, reflective thinking), aspects which are all more closely related to cognition than to affect.

With regard to *attitudes towards mathematics*, their transcendence in the process of teaching-learning and in students' mathematical performance is well known (Miñano & Castejón, 2011; Miranda, 2012; Sakiz, Pape, & Hoy, 2012). The influence of positive attitudes towards mathematics on anxiety is also well established (Akin & Kurbanoglu, 2011). In this regard, some works have found that students with better attitudes towards mathematics have higher perceptions of the utility of mathematics, denoting intrinsic motivation towards their study (Perry, 2011), they have a better mathematical self-concept (Hidalgo, Maroto, & Palacios, 2005), are more confident they can learn mathematics (McLeod, 1992) and, especially, they display approach behaviors towards mathematics (Fennema & Sherman, 1976).

Due to their importance, attempts to measure attitudes towards mathematics appear early on, and the works of Aiken (Aiken, 1972, 1974, 1979; Aiken & Dreger, 1961), with the contributions of Dutton and Blum (1968), are pioneer in this topic.

In one of the first measurement instruments of these attitudes, Aiken and Dreger (1961) prepared a questionnaire made up of 20 items with two subscales: *Pleasure* and *Fear of Mathematics*. As these dimensions can be considered the extreme poles of the same continuum, some authors have considered it a unidimensional scale (Auzmendi, 1992). In a later version, Aiken (1972) introduced the factor *Enjoyment of Mathematics*. Two years later, the same author (Aiken, 1974) presented what is no doubt one of the most frequently used scales in the measure of attitudes towards mathematics, comprised of two subscales: the *Value of Mathematics* scale and the *Enjoyment of Mathematics* scale. In a later version, Aiken (1979) increased the number of factors to a total of four: *Enjoyment of Mathematics*, *Mathematical Motivation*, *Value-Utility of Mathematics*, and *Fear of Mathematics*.

There have been numerous adaptations of these scales (Aiken, 1974, 1979), coinciding with the original reliability values and with the factor structure of the four subscales.

The scale of Fennema and Sherman (1976) is, in the words of Tapia and Marsh (2004), the most popular measure of attitudes towards mathematics of the last three decades. The origin of this scale lies in the study of differences between men and women in their attitudes towards mathematics as well as their influence on performance. This scale has been the object of extensive studies and it has been translated into various languages, and modified for application in different situations.

The contribution of Tapia and Marsh (2004), The *Attitude toward Mathematics Inventory* (ATMI), is doubtless one of the most extensively used instruments to measure attitudes towards mathematics. Its final version is made up of 49 items that attempt to assess six aspects of these attitudes: *Confidence-Self-concept*, *Anxiety*, and *Utility-Value of Mathematics*, *Enjoyment of Mathematics*, *Motivation* and *Parents' and Teachers' Expectations*.

Among the most recent contributions in English, we underline the work of Kadijevich (2008), from the TIMSS-2003 report, as well as those of Tahara, Ismailb, Zamanic, and Adnand (2010). Adelson and McCoach (2011) prepared a scale of attitudes towards mathematics for primary education students, which they called *The Math and Me Survey* and which, after the preliminary analyses, presented two factors related to the perception of efficacy and enjoyment of mathematics.

The adaptations to Spanish of the scales of Aiken (1974) and of Fennema and Sherman (1976), as well as the later ones of Tapia and Marsh (2004) are scarce and generally not oriented toward psychometric analysis. Such is the case of the adaptation of Cazorla, Silva, Vendramini, and Brito (1999) of the scale of Aiken (1974) based on a prior Portuguese scale of Brito (1998), for the study attitudes towards statistics; the scales of Quiles (1993), which attempt to relate attitudes towards mathematics and academic performance; the more modern scale of Estrada and Díez-Palomar (2011), focused on the mathematical education of relatives; or the scale of González-Pienda, Fernández-Cueli, García, Suárez, Fernández, Tuero-Herrero, and Helena da Silva (2012), aimed at determining differences in the mathematical attitudes of men and women. This lack of adaptations to Spanish was noted in the early works of Gairín (1990) and more recently by Muñoz and Mato (2008, who also mentioned the inexistence of adaptations of these scales in our context.

As mentioned, the work of Gairín (1990) can be considered pioneer in the measure of attitudes towards mathematics in the Spanish language. In this work, the author mentions the need for an instrument to measure attitudes towards mathematics in Spanish because, at that time, all the known scales originally proceeded from the Anglo Saxon world. Among the author's proposals is his verbal scale made up of 22 items rated on a Likert scale, with three dimensions related to enjoyment of mathematics, utility of mathematics, and confidence-anxiety towards mathematics. The reliability indexes of these three factors, obtained with the test-retest technique, showed correlations ranging from .77 to .93 in the time interval, and the reliability of the entire scale was .84.

Auzmendi (1992) designed what is clearly the scale of attitudes towards mathematics that is the most cited of those created in Spanish. As in Gairín (1990), the author justifies the elaboration of a new scale on the basis of the lack of this kind of instruments in Spanish. The final test has 25 items that, after the corresponding factor analyses, present five main components: *Feelings of Anxiety and Fear* towards mathematics manifested by the student, *Liking-Enjoyment of Mathematics*, *Utility of Mathematics*, and *Motivation* and *Confidence*. Cronbach's alpha of these scales ranges between .91 for the Anxiety scale, and the lower value of .49 for the Confidence scale. The validation sample was made up of 1,221 secondary and high school students.

As instruments equally distant in terms of time, we note the contributions of Escudero and Vallejo (1999) who prepared an instrument to measure attitudes towards mathematics using a total of 18 items related to enjoyment, utility, and motivation. One year before, Bazán and Sotero (1998) had prepared the "Escala de Actitudes hacia las Matemáticas" (in English, Attitudes towards Mathematics Scale; EAHM-V), made up of 31 items divided into four dimensions: *Affect*, *Applicability*, *Reliability*, and *Anxiety*. The scale was aimed at measuring the attitudes of students who had just entered the university, and obtained a reliability of .90 for the total scale in a sample of 256 university students.

In recent years, new proposals have been made, among which are notable the contributions of Muñoz and Mato (2008) and of Alemany and Lara (2010). Muñoz and Mato (2008) presented a scale of attitudes towards mathematics, designed using a sample of 1,220 secondary education students. The final questionnaire had 19 items that, after the corresponding factor analysis, presented two factors: *Teacher's Attitude as perceived by the student* and *Enjoyment-Utility of Mathematics*. The final version had a reliability of .97. Lastly, we note the contribution of Alemany and Lara (2010), who designed and validated a new scale of attitudes towards mathematics for secondary students, made up of 37 items. A differentiating element of this work is that the validation sample was made up of students of Berber ethnic origin and it was prepared both in Spanish and in the Tamazight language. The final version obtained a Cronbach's alpha of .92 in a sample of 236 students from second and third grade of secondary education.

In summary, the different scales of attitudes towards mathematics, both in English and in Spanish, generally present adequate reliability indexes if the limitations of the Cronbach alpha coefficient to assess reliability are not taken into account (see our comment in the section 'Evidence of reliability and internal consistency'). Nevertheless, the disparity of the subscales hinders reaching a coherent—or at least a unified—interpretation of the construct of attitudes towards mathematics. In a large number of these scales, at least in the Spanish versions, the psychometric values were obtained from small samples, and mostly from students in Compulsory Secondary Education.

In the following paragraphs, we will present a multidimensional scale of attitudes towards mathematics with items taken from the corresponding references adapted to our current historical-cultural setting, with a very large sample of primary, secondary, and high school students, and with psychometric values obtained from the proposals of the classical test theory, the structural equation models (measurement model), and the item response theory (Hambleton, Swaminathan, & Rogers, 1991; Samejima, 1969, 2010).

Method

Participants

The study was carried out with a sample of 4,807 students from 14 public, private, and subsidized schools and institutes from Spanish provinces. Of the 14 participant centers, 3 were from Segovia (25%), 3 from Ávila (8% of the student body), 3 from Soria (10%), 2 from Valladolid (40%), and 3 from Zamora (17%). Two of these centers were private and/or subsidized schools or institutes. The data from these centers was collected in all the trajectories of the selected courses. Of the participants, 67% studied in schools located in province capitals, and the remaining 34% in rural areas. The participant schools were selected by means of stratified random sampling, taking the geographical area, the educational level, and the official status of the center as selection stratum. Participants' mean age was 14 years, ranging from 11 to 23 years. Distribution by educational levels is shown in Table 1. Of the participants, 53% were male and 47% were female. Their grades in mathematics had a normal distribution with a mean value of 5.62 (SD = 1.95).

Table 1

Participants' Educational Level

		N	%	% males	% females
Valid	6th. Primary	394	8.20	53.8	46.2
	1st. CSE	828	17.22	54.0	46.0
	2nd. CSE	1,035	21.53	55.1	44.9
	3rd. CSE	1,267	26.36	52.2	47.8
	4th. CSE	680	14.15	49.5	50.5
	1st. HE	348	7.24	51.4	48.6
	2nd. HE	189	3.93	51.3	48.7
	Total	4,741	98.63	52.8	47.2
Missing		66	1.37	—	—
Total		4,807	100.00	53.0	47.0

Variables and instruments

To construct the *Escala de Actitudes hacia las Matemáticas* (EAM; in English, the Scale of Attitudes towards Mathematics), we drew from the works reviewed in the previous section that present five generalized factors: liking-enjoyment of mathematics, anxiety towards mathematics, perception of difficulty, perceived utility, and mathematical self-concept were the thematic fields chosen to prepare the initial items of the test.

First, a broad set of questions related to these five factors was designed. To assess the factors associated with *liking or enjoyment of mathematics* the *Enjoyment of Mathematics* subscale of Aiken (1974) and the *Liking scale* of Fennema and Sherman (1976) was used. To select the questions related to *anxiety towards mathematics*, we drew from the works of Richardson and Suinn (1972). To measure the *perception of difficulty of mathematics*, we drew from the previously cited works of Aiken (1974) and Fennema and Sherman (1976). The questions of the factor of *perception of utility of mathematics* were developed based on the proposals of Aiken (1974) and Fennema and Sherman (1976). The design of the questions related to the *perception of efficacy and/or competence in mathematics* (mathematical self-concept) was based on prior works in this type of measurement instruments such as those of Pietsch, Walker, and Chapman (2003).

All these questions were assessed by experts in Didactics of Mathematics. Through these assessments, the most pertinent questions due to their relevance (the items must be clearly related to the object of study) and clarity (simple, easily understood statements) were selected. A pilot study on a small sample with this selection was carried out. After the elimination and/or selection of the most adequate items, we prepared the final scale made up of a total of 37 questions, which are presented grouped by factors in Table 3. In this final scale, all the items are rated according to the degree of agreement with the statement on a 5-point Likert-type scale (values ranging from 0 to 4).

Procedure

The scales were administered by the authors and collaborating teachers during the academic courses 2009/2010, 2010/2011 and 2011/2012. The scales were anonymous and completed by the subjects of the sample in the presence of the teacher and/or collaborator. Before collecting the data, both the parents' informed consent and the authorization of the headmasters of the schools were obtained.

Results

Exploratory factor analysis

The original sample (n = 4,741) was divided into two randomly extracted subsamples ($n1$ = 2,371 and $n2$ = 2,370). The first one was used to carry out the exploratory factor analysis (EFA) and the second was used as a validation sample for the confirmatory factor analysis (CFA) and for the analyses based on the item response theory, described below. Three items were eliminated ("I am happier about getting a 10 in mathematics than in any other subject", "My parents are more concerned about my outcomes and grades" and "When I have trouble with mathematics, I usually ask for help from my family") because they presented corrected homogeneity indexes lower than .20. The distributions of the variables age and sex were similar in both subsamples. The standardized Pearson residuals ranged from –1.09 to 1.03, and the model [AGE, SAMPLE][SEX] was nonsignificant ($\chi^2_{(15)}$ = 11.859, p = .690), so the hypothesis of equivalent subsamples was accepted.

Exploratory factor analysis was carried out with the SAS, v. 9.2 program. When determining the factor structure of the EAM, we used two extraction procedures (*Principal Axis Factoring*, PAF, and *Maximum Likelihood*, ML) to verify whether both methods obtained comparable results. Both analyses were carried out on the polychoric correlation matrixes, in view of the ordinal nature of the input data. The adequacy of the input data was confirmed by means of Bartlett's sphericity test, the KMO index, and the matrix determinant (Table 3).

Both the pattern matrix and the structure matrix obtained similar results in the first subsample; the items presented a practically identical loading distribution in the different factors. This similarity was corroborated by means of the Pearson correlations and the congruence coefficients. Table 2 shows that the Pearson correlations reached a mean of .991, ranging from .982 to .996. The congruence coefficients calculated from the pattern matrix loadings ranged from .987 to .997 (exceeding the limit of .95, habitually considered acceptable for this type of analysis). Similar results were obtained when comparing the matrixes of the two random subsamples using the PAF extraction method. The Pearson correlation coefficients ranged from .901 to .986, and the congruence coefficients from .955 to .997.

PROMAX oblique rotation was used, because prior research indicated that the dimensions of anxiety towards mathematics are correlated (e.g., Pajares & Miller, 1994).

Table 2

Pearson Correlations and Congruence Coefficients

	Extraction methods		Random subsamples	
	r	CC	*r*	CC
F1	.994	.997	.986	.997
F2	.996	.997	.980	.994
F3	.992	.994	.962	.985
F4	.982	.987	.901	.955

The results of the EFA presented in the following paragraphs correspond to the PAF extraction method, in view of the more 'classical' nature of this method compared to the ML method (Pett, Lackey, & Sullivan, 2003, p. 103).

To determine the number of factors to retain, various criteria into account were taken: the Kaiser-Gutman rule, Cattell's scree-test and parallel analysis. The Kaiser-Guttman rule (Eigenvalues higher than 1.00) suggested retaining five factors, as did the results of the scree-test. However, we decided to disregard this recommendation because both methods usually lead to overfactorization; in fact, the fifth factor presented an Eigenvalue of only 1.10, it explains less than 3% of the common variance, and only includes two items in a dimension that could be called *Learned helplessness* ("Except for a few cases, no matter how much effort I put out, I cannot understand mathematics" and "No matter what I do, I always get low grades in mathematics"). As the five- factor structure is not statistically justifiable, an optimized parallel analysis (Timmerman & Lorenzo-Seva, 2011) was conducted by comparing the Eigenvalues obtained through analysis with the randomly generated Eigenvalues of 1000 subsamples obtained from the original sample. This analysis is currently considered the most adequate to make decisions about the number of factors to retain (Hayton, Allen, & Scarpello, 2004). As of the fourth factor, the magnitude of the randomly generated Eigenvalues exceeded that of the Eigenvalues obtained through the analysis, so we decided to retain the four-factor solution.

In accordance with the conventional criteria in this type of analysis, there were three item-retention criteria: (a) the item loading on the main factor should at least reach the value of .40; (b) the loading on the remaining factors should not exceed the value of .35; and (c) the

Table 3

Pattern Coefficients, Structure Coefficients, and Communalities

		Pattern Coefficients				Structure Coefficients				h2
		F1	F2	F3	F4	F1	F2	F3	F4	
i29	In maths it's hard for me to decide what I have to do	.750	–.048	.064	.067	.769	.407	.327	.297	.598
i28	Usually I feel unable to solve mathematical problems	.738	.025	.019	.140	.803	.480	.324	.386	.665
i36	Usually I have difficulty with mathematics	.732	.148	–.128	.175	.820	.544	.238	.439	.727
i34	I feel more awkward in maths than most of other students	.722	–.084	.005	.202	.743	.379	.265	.400	.587
i35	Mathematics confuse me	.700	.112	.063	.056	.800	.534	.379	.329	.661
i22	I have always had trouble with maths	.691	–.028	.001	.230	.749	.427	.281	.438	.606
i25	No matter what I do, I always get low grades in mathematics	.687	–.085	–.072	.024	.623	.257	.143	.196	.400
i10	In mathematics my mind often blanked out on me	.660	.146	.014	–.040	.730	.489	.315	.225	.549
i27	I do not know how to study mathematics	.651	.104	.048	.024	.732	.482	.337	.277	.549
i14	Except for a few cases, no matter how much effort I put out, I cannot understand mathematics	.642	–.079	–.025	–.039	.578	.237	.166	.130	.343
i23	I have no idea what mathematics are	.533	–.186	.414	.145	.631	.344	.548	.317	.545
i12	I will always have difficulty on learning mathematics	.457	.082	.079	.046	.544	.379	.291	.235	.313
i32	I'm one of those people who were not born to learn maths	.456	.012	.202	.028	.544	.359	.378	.212	.335
i05	When I study mathematics, I feel more uncomfortable than when I do it with other subjects	.391	.409	.040	–.130	.583	.587	.348	.154	.462
i08	I enjoy studying maths	.018	.848	–.022	.073	.485	.875	.390	.392	.771
i39	When I have to do math homework, I do it with some joy	–.059	.835	–.081	.112	.392	.808	.303	.391	.671

		Pattern Coefficients				Structure Coefficients				h2
		F1	F2	F3	F4	F1	F2	F3	F4	
i38	I can spend hours studying and doing math problems, time goes by so fast!	–.089	.785	–.140	.107	.312	.713	.210	.348	.540
i20	If given the opportunity, I would choose elective courses related to mathematics	–.223	.690	.018	.245	.230	.672	.300	.436	.525
i16	The subject taught in mathematics classes is very interesting	–.086	.685	.169	.153	.389	.774	.481	.412	.641
i19	Mathematics is one of the most boring subjects	.257	.671	.157	–.221	.602	.798	.522	.139	.741
i01	I like mathematics	.134	.669	.063	.199	.576	.843	.456	.502	.769
i31	Studying mathematics is dead boring	.222	.660	.184	–.159	.591	.804	.542	.190	.731
i02	I feel comfortable doing math problems	.158	.610	–.043	.239	.543	.764	.339	.510	.659
i11	It's time for maths, how awful!	.291	.605	.188	–.204	.618	.771	.538	.147	.717
i17	I hate studying maths, even the easiest parts	.293	.481	.228	–.175	.577	.677	.526	.137	.591
i09	Mathematics are easy	.310	.432	–.130	.286	.583	.644	.233	.522	.578
i06	Mathematics are useless	.032	.071	.781	.022	.361	.457	.829	.196	.696
i15	Mathematics are useful and nececssary in all areas of life	–.183	.088	.699	.282	.208	.419	.723	.381	.611
i07	Mathematics should be present only in science careers	.162	–.007	.673	–.063	.384	.367	.718	.104	.537
i21	Learning math is a matter of a few	.378	–.115	.538	–.006	.511	.333	.622	.165	.488
i04	I want to learn maths	–.292	.411	.436	.315	.185	.575	.575	.454	.576
i30	I can become a good student of mathematics	.250	–.021	.250	.615	.524	.457	.440	.730	.680
i13	If I set my mind I it would come to dominate mathematics well	.138	–.094	.310	.589	.387	.343	.421	.652	.536
i37	I am good at mental calculation	.143	.127	–.119	.525	.333	.345	.085	.597	.395
i33	I am good at mathematics	.442	.280	–.163	.465	.679	.614	.210	.680	.754

		Pattern Coefficients				Structure Coefficients				h2
		F1	F2	F3	F4	F1	F2	F3	F4	
i18	For my professional future mathematics is one of the most important subjects	–.142	.247	.374	.426	.261	.504	.512	.540	.509
i26	For my math teachers and lecturers I am a good student	.277	.201	–.068	.399	.485	.466	.196	.550	.434
	Eigenvalue*	13.810	2.398	1.551	1.333					
	Proportion of Common Variance Explained*	.432	.075	.049	.042					

Correlation matrix adequacy*	Bartlett's Sphericity Test	$\chi^2_{(666)} = 7948.11, p = .000$
	Kaiser-Meyer-Olkin Index	$KMO = .972$
	Matrix Determinant	.000

Note. Shown in gray the items removed from the final version.

* Values obtained with the final version of 32 items.

difference between the loading on the main factor and the loadings on the remaining factors should exceed .15.

As seen in Table 3, all the items met the first criterion. The second criterion was not met by Items 5 ("When I study mathematics I am more uncomfortable than with other subjects"), 23 ("I have no idea what mathematics are about"), 4 ("I want to learn mathematics"), and 33 ("I'm good at mathematics"). These items presented crossloadings, so it seems that they could be confusing for the students, or their content could not be clearly ascribed to the factors considered. In addition to these items, Item 18 ("Mathematics is one of the most important subjects for my professional future") did not meet the third criterion. Thus, these five items were eliminated from further analysis, and the final version of the scale was made up of 32 items.

The first factor (*Perception of Mathematical Incompetence*) is made up of 12 items and explains 43.2% of the common variance. It includes items related to the perception of inability, awkwardness, confusion, difficulty and expectations of failure. This factor was present in the first works of Fennema and Sherman (1976) and Sandman (1980) as well as in the more recent ones of Kadijevich (2008), Tahara et al. (2010) and Adelson and McCoach (2011), with a marked negative valence in the attitude towards mathematics.

The second factor (*Enjoyment of Mathematics*) is made up of 12 items and explains 7.5% of the common variance. The items refer to the positive emotions evoked by the study of mathematics, perception of ease and comfort when solving mathematical problems. As before, this factor was present in the first scales of attitudes (Aiken & Dreger, 1961; Aiken, 1972, 1974, 1979) as well as in the more modern ones of Tapia and March (2004), Adelson and McCoach (2011), and Muñoz and Mato (2008). In all cases, the positive nature of the factor, associated with enjoyment of mathematics and liking its study, is mentioned.

The third factor (*Perception of Utility*) is made up of 4 items and explains 4.9% of the common variance. The content of the items refers to the utility of and need for mathematics. This same factor was found in the contributions of Fennema and Sherman (1976), Aiken (1972, 1974, 1979), Tapia and March (2004), Sandman (1980), Tahara et al. (2010), Kadijevich (2008), Adelson and McCoach (2011) and Auzmendi (1992).

The fourth factor (*Mathematical Self-concept*) includes 4 items and explains 4.2% of the common variance. The items refer to the student's self-concept as being skilled and capable of studying mathematics. As a specific factor of the scales of attitudes towards mathematics, it can be found in the works of Fennema and Sherman (1976), Tapia and Marsh (2004), and Alemany and Lara (2010).

Evidence of reliability and internal consistency

Although Cronbach's alpha coefficient is historically the most used in the literature on psychological research to assess reliability, it has recently been seriously questioned, as it is not related to the internal structure of the test, given the item covariance matrix and the habitual assumptions about measurement errors, and it cannot be stated that alpha measures internal consistency or unidimensionalality (e.g., see Sijtsma, 2009). An alternative is the calculation of composite reliability based on the loadings and measurement errors (this coefficient is provided in the section on CFA), the ordinal alpha coefficient (if a factor analysis model is assumed) or the ordinal theta coefficient (if a principal component analysis model is assumed), the Omega coefficient of McDonald, and the greatest lower bound (glb).

Therefore, internal consistency was estimated by calculating the ordinal modality of the Cronbach's alpha and theta coefficients (Zumbo, Gadermann, & Zeisser, 2007) of the subscales (with the corresponding 95% confidence intervals). To determine the internal consistency of the global scale, we calculated the stratified alpha coefficient (to correct alpha's underestimation of internal consistency when a scale contains intercorrelated factors), as well as the Ω and glb coefficients. The Ω coefficient can be interpreted as the squared correlation between the scale score and the latent variable common to all the indicators, or 'universe' of indictors, of which the items of the scale are a subset (McDonald, 1999), while glb represents the smallest possible reliability given the observed covariance matrix of the items, under the restriction that the sum of the error variances is maximized for the errors that present $r = 0$ with the rest of the variables (Ten Berge, Snijders, & Zegers, 1981). This information is presented in Table 4, which shows sufficient evidence of reliability for the individual factors and the global scale.

Table 4

Internal Consistency of the Scale of Attitudes Towards Mathematics

Factor	α_{ORD}	CI 95%	Θ	Inter-item Correlations M	Inter-item Correlations SD
Perception of Mathematical Incompetence	.887	[.882, .892]	.918	.440	.121
Enjoyment of Mathematics	.921	[.918, .924]	.941	.492	.102
Perception of Utility	.678	[.663, .693]	.779	.354	.063
Mathematical Self-concept	.679	[.664, .694]	.737	.346	.123
Total Scale	.933	[.930, .936]	.958	.438	.120
Stratified α			.949		
Ω			.939		
glb			.965		

Note. α_{ORD} = Ordinal Cronbach's Alpha; Ω = McDonald's Omega; glb = greatest lower bound to reliability.

Confirmatory factor analysis

On the basis of the results obtained in the exploratory factor analysis, we conducted CFA on the second subsample (N = 2370).

The analysis was carried out on the 32 items retained according to the results of the EFA. The following six models were tested: unifactorial (M1), four correlated factors (M2), hierarchical, with four first-order factors and one second-order factor (M3), three correlated factors (M4), exploratory structural equation modeling (ESEM; M5), and bifactor model (M6). The ESEM models are a recent derivation of the measurement models, estimated with CFA but without the restriction that the factors on which an item does not load must have a regression coefficient equal to zero. In this sense, they are an integration of the exploratory and confirmatory models (Morin, Marsh, & Nagengast, 2013). The bifactor models include regression coefficients of the items from the individual factors and from a general factor (e.g., see Reise, Morizot, & Hays, 2007). The BI-GEOMIN rotation method—oblique rotation in which the specific factors are correlated with each other and with the general factor (Muthén & Muthén, 2013) was used.

Two programs were used to estimate the parameters: LISREL, v. 9.1 for the first four models and MPlus v. 7.1 (Muthén & Muthén, 2013) for the last two because the current configuration of LISREL does not allow the estimation of ESEM or bifactor models.

In models M1, M2, M3, and M4, the asymptotic covariance and variance-covariance matrixes[1] were used as input data and, as estimation method, DWLS due to the ordinal nature of the input data (Edwards, Wirth, Houts, & Xi, 2012; Hu & Bentler, 1999). Models M5 and M6 were estimated using the polychoric correlations matrix with the robust maximum likelihood method. For M5 (ESEM), the GEOMIN oblique rotation method was used. In Table 5 the fit indexes according to the different models are presented.

As seen in Table 5, the models with the best fit were M2 (four correlated factors) and M5 (ESEM). The root mean square error of approximation (RMSEA) value was adequate in both models, as was standardized root mean square residual (SRMR). However, the CFI and TLI fit indexes were better in the M2 model, so this model was chosen for the final version of the scale. It was also checked the possible presence of a method effect associated with the negatively worded items (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003). It was ruled out that these items share an unexplained common variance or a variance related to their respective latent variables, but associated with their negative wording.

Table 5

Fit Indexes of Measurement Models

	M1	M2	M3	M4	M5	M6
GL	464	458	460	461	374	432
χ^2	6354.73	3199.45	4047.43	4446.50	2074.45	3356.75
p	.000	.000	.000	.000	.000	.000
RMSEA	.053	.047	.084	.060	.045	.055
(RMSEA 90% CI)	(.051; .054)	(.045; .048)	(.082; .085)	(.059; .062)	(.043; .047)	(.053; .057)
CFI	.970	.986	.982	.979	.950	.914
TLI	.968	.985	.980	.978	.934	.902
SRMR	.078	.051	.115	.055	.023	.047

Note. RMSEA: Root Mean Square Error of Approximation; CFI: Comparative Fit Index; TLI: Tucker-Lewis Index; SRMR: Standardized Root Mean Square Residual.

All the nonstandardized regression coefficients were statistically significant, with *t* values ranging from 17.54 to 73.10. The standardized coefficients ranged from .3 to .9: the smallest (λ_x = .498) corresponded to Item 37 ("I am good at mental calculation") and the largest (λ_x = .879) corresponded to Item 1 ("I like mathematics").

Table 6 presents the values of average variance extracted (AVE), construct reliability, and McDonald's omega coefficients corresponding to the four correlated factors model. The factor *Mathematical Self-concept* obtained the lowest standardized regression coefficients (*M* = .630) and the *Enjoyment of Mathematics* factor obtained the highest (*M* = .753). The intermediate magnitudes corresponded to the factors *Perception of Mathematical Incompetence* (*M* = .734) and *Perception of Utility* (*M* = .681). It is consequently concluded that the scale provides sufficient evidence of reliability.

Table 6

Average Variance Extracted, Construct Reliability, and McDonald's Omega (four correlated-factor solution)

	F1	F2	F3	F4
Average Variance Extracted	.547	.623	.569	.474
Construct Reliability	.934	.952	.804	.770
McDonald's Omega	.932	.951	.798	.749

Calibration of the EAM

We conducted an analysis of the EAM by means of the graded response model (GRM, Samejima, 1969, 2010), after performing the necessary confirmation of unidimensionality and local independence of each one of the four subscales. Summing up, the results revealed that: (a) all the α_i discrimination parameters were adequate according to the classification of Baker (2001), including a range of 1.06 to 3.36; 6 of them were moderate (from 0.65 to 1.34), 5 were high (from 1.35 to 1.69), and 21 were very high (higher than 1.7); (b) the standard errors of the α_i parameters were very low, with a range of .04 to .13; (c) the order of all the β_{ik} localization parameters corresponded with our expectations for the model because the thresholds were never disordered; (d) the ranges of the β_{ik} parameters were similar in the four subscales (4.41 in *Perception of Mathematical Incompetence*, 4.39 in *Enjoyment of Mathematics*, 4.33 in *Perception of Utility* and 5.10 in *Mathematical Self-concept*), showing that they cover an extensive range of the latent variables measured; (e) the range of the standard errors of the β_{ik} parameters was always very low: from .02 to .09 for the first factor, from .02 to .06 for the second, from .01 to .12 for the third, and from .02 to .08 for the fourth; (f) all four dimensions presented an adequate global fit, as convergence was reached at less than 50 iterations in all cases, the standard errors and the M_2 values were reduced and the RMSEA values were lower than .06; (g) analysis of the individual fit of the items, calculated by means of the SAS macro IRTFIT (Bjorner, Smith, Stone, & Sun, 2007) produced G^2 and χ^2 values with p-values higher than .05 in all items; (h) the invariance of the parameters with two randomly extracted subsamples (*n1* = 1185, *n2* = 1185) were confirmed after rerunning the complete estimation process of the α_i and β_{ik} parameters on each sample; and (i) none of the items presented a uniform differential item functioning (DIF) according to the results of the contrast method of Benjamini-Hochberg (D. Thissen, personal communication, January 20, 2012). These results provide solid support for the adequacy of the EAM as a measure of attitudes towards mathematics.

Validity evidence

The evidence of *content validity* of the EAM, understood as the correspondence between the sample of indicators used and the domain one wishes to measure, as well as the representativeness, relevance, and technical quality of the items, is partially supported by the bibliographic review carried out (the individual items and the domains or dimensions both correspond, to a great extent, to those found in prior research) and by the discriminatory power of the items, as confirmed by the α_i parameters in the analysis of the GRM model.

The evidence of *construct validity* is sufficiently supported by the results of the analyses of the internal structure of the EAM, explained in the sections of confirmatory factor analysis and calibration of the scale through GRM.

The evidence of *convergent validity* of the measurement model can be established by means of the magnitude, direction, and statistical significance of the regression coefficients (Hair, Black, Babin, & Anderson, 2010). Firstly, all the coefficients were significantly different from zero, as indicated by the associated t values (all higher than 3.29, denoting a value of $p < .001$). Secondly, they were all higher than .50 (20 of them were higher than .70, and none exceeded the value of .90, which could indicate the presence of multicolinearity). Thirdly, all the coefficients were positive, according to our expectations for the model. Fourthly, the values of AVE (see Table 6) were acceptable in Factors 1 (*Perception of Mathematical Incompetence*), 2 (*Enjoyment of Mathematics*), and 3 (*Perception of Utility*) as well as in the total scale. The value obtained for Factor 4 (*Mathematical Self-concept*) was somewhat questionable, although it approached the recommended value of .50. Lastly, in all cases, the recommended value of .70 for the composite reliability and McDonald's omega coefficients for the assessment of the construct reliability in the first case and of the general loading of the scale in the second (McDonald, 1999) was amply exceeded. Taken conjointly, these results allow us to state that the evidence of convergent validity of the scale is sufficient.

Finally, to determine the evidence of *discriminant validity* (showing that the measure in each of the constructs assessed is different from the rest), the estimations of the AVE of each factor were compared with the associated squared interconstruct correlations. As can be observed in Table 7, except for that of the factor *Mathematical Self-concept*—which was only slightly lower than the squared correlation between Factors 1 and 4—all the AVE estimations were higher than the squared correlations between the factors, providing evidence of the discriminant validity of the instrument.

Discussion

Attitudes towards mathematics are currently an area of great value for research within what is known as the affective domain in mathematics. Their importance is obvious due to the large number of studies dedicated to their conceptualization and measurement and to the range of topics to which they have been related.

Table 7

Evidences of Sub-Scales Discriminant Validity

Subscale	F1	F2	F3	F4
Perception of Mathematical Incompetence	1.000	.530	.312	.484
Enjoyment of Mathematics	.728	1.000	.442	.433
Perception of Utility	.559	.665	1.000	.407
Mathematical Self-concept	.696	.658	.638	1.000

Note. Squared inter-factor correlations above diagonal. Inter-factor correlations below diagonal. Standardized variances on diagonal.

Although, as a rule, there has been much disparity in the attempts to measure these attitudes, some important common points have also emerged, especially concerning the factor structure of the construct attitudes towards mathematics. Specifically, liking or enjoyment of mathematics, the value-utility granted to the discipline, perception of self-efficacy, and mathematical anxiety have been present in an important part of the research on this topic.

This study had the goal of contributing to the clarification of the internal structure of the construct 'attitudes towards mathematics' by elaborating a measurement scale with solid psychometric properties. To our knowledge, this is the first instrument that was analyzed and calibrated in two very large random subsamples (more than 2,000 participants each one), extracted in turn from a probabilistic sample. It is also among the very few investigations that have used models from the framework of the item response theory to calibrate the measuring instrument and which provides evidence of statistical reliability based on the ordinal nature of the data. The scale has shown sufficient evidence of validity and reliability. Such evidence, along with the guarantee of having been obtained from a very large sample (with the guarantees of statistical power and decrease of measurement error) in which all the non-university educational levels are represented, allows us to conclude that this Scale of Attitudes towards Mathematics is solid and robust, and is of great potential utility for non-university educational levels.

Attitudes towards mathematics, especially the self-perception of mathematical competence and enjoyment of mathematics, are no doubt areas of great interest for researchers in education, educational psychologists, and school counselors, especially at times of transition (e.g., from Primary to Compulsory Secondary Education or from High School to University). In this sense, the EAM could contribute to furthering our understanding of the conceptual and empirical bases of attitudes towards mathematics in young students, before their admittance to university.

It is quick and easy to administer and allows identification of students with low scores on the scale (especially worrisome are students who obtain low scores on the factors Mathematical Incompetence and Mathematical Self-concept) and designing of actions to improve both attitudes and competence. It could also be useful in the case of students with sufficient mathematical skills but who show poor attitudes, perhaps due to problems with self-concept or because they compare themselves with more skilled students. It could also detect students with poor attitudes who are at risk of behavior problems in the mathematics class. Lastly, it could be interesting for teachers to assess the initial attitudes (at the beginning of the school term) and to confirm whether they vary (in successive applications of the EAM) over the term, that is, whether the school setting has a positive or negative effect on these students' self-perception and enjoyment of mathematics.

This study also presents some weak points that could, in turn, become future pathways for research in this field.

For example, to avoid overextending this work, in this study, the reliability of the scale was limited to analysis of internal consistency and the standard error of measurement. In the future, it should be complemented with other evidence of reliability, especially with regard to the temporal stability of the scores.

Concerning validity, other evidence of validity, especially external, consequential, and nomological validity, should be investigated. The evidence of external validity should be studied in the dual sense of generalizability (the extent to which the scores and interpretation can be generalized to groups of populations, situations, and tasks) and the relation between attitudes and performance and interest in mathematics (e.g., how self-perceptions, enjoyment of mathematics, and the perception of its utility influence performance and the choice of studies based on mathematics). It is also recommended to study the validity of the instrument with reference to consequential aspects (e.g., evidence-based potential and real consequences of using the scale) (Messick, 1995), and its nomological validity (the extent to which the EAM predicts other concepts from the theoretical model on which it is based).

It would also be interesting to investigate the possible variations of attitudes towards mathematics (e.g., by means of the analysis of structured means modeling) as a function of sex, intellectual level, ethnic group, educational level, teachers' evaluation, or students' performance in mathematics and in other disciplines of the curriculum, as well as to examine whether such attitudes are stable over time or whether they can be changed at the mid or long term by implementing programs of sensitization and improvement.

Lastly, the fact that internal consistency assessed by means of the ordinal alpha coefficient was found to be higher in the factor Enjoyment of Mathematics than in the remaining factors poses the question of the extent to which contextual factors (e.g., teachers' instructional competence, the design of the curriculum, or the teaching methods) may affect students' attitudes towards mathematics. Research of these aspects could be an important contribution to improve teaching practices of this discipline.

References

Adelson, J. L., & McCoach, D. B. (2011). Development and psychometric properties of the math and me survey: Measuring third through sixth graders' attitudes toward mathematics. *Measurement and Evaluation in Counseling and Development, 44*(4), 225-247. doi: 10.1177/0748175611418522

Aiken, L. R. (1972). Research on attitudes toward mathematics. *The Arithmetic Teacher*, *19*(3), 229-234.

Aiken, L. R. (1974). Two scales of attitude toward mathematics. *Journal for Research in Mathematics Education*, *5*, 67-71.doi: 10.2307/748616

Aiken, L. R. (1979). Attitudes toward mathematics and science in Iranian middle schools. *School Science and Mathematics*, *79*, 229-234. doi: 10.1111/j.1949-8594.1979.tb09490.x

Aiken, L. R., & Dreger, R. M. (1961). The effect of attitude on performance in mathematics. *Journal of Educational Psychology*, *52*, 19-24. doi: 10.1037/h0041309

Akin, A., & Kurbanoglu, N. (2011). The relationships between math anxiety, math attitude and self-efficacy: A structural equation model. *Studia Psychologica*, *53*(3), 263-273.

Alemany, I., & Lara, A. I. (2010). Las actitudes hacia las matemáticas en el alumnado de la ESO: un instrumento para su medición. *Publicaciones*, *40*, 49-71.

Auzmendi, E. (1992). *Las actitudes hacia la matemática-estadística en las enseñanzas medias y universitarias. Características y medición*. Bilbao: Mensajero.

Baker, F. (2001). *The basics of item response theory*. College Park: ERIC

Clearinghouse on Assessment and Evaluation, University of Maryland.

Bazán, J., & Sotero, H. (1998). Una aplicación al estudio de actitudes hacia la matemática en la Unalm. *Anales Científicos UNALM*, *36*, 60-72.

Bjorner, J. B., Smith, K. J., Stone, C., & Sun, X. (2007). *IRTFIT: A Macro for Item Fit and Local Dependence Tests under IRT Models*. Quality Metric Incorporated. University of Pittsburgh: School of Education.

Brito, M. R. F. (1998). Adaptação e validação de uma escala de attitudes em relação à matemática. *Zetetiké, 6*(9), 109-162.

Cazorla, I. M., Silva, C. B., Vendramini, C., & Brito, M. R. F. (1999). Adaptação e validação de uma escala de attitudes em relação à estatística. Actas de la Conferência Internacional *Experiências e Perspectivas do Ensino da Estatística. PRESTA,* Florianópolis: Florianópolis.

Dutton, W. H., & Blum, M. P. (1968). The measurement of attitudes toward arithmetic with a Likert-type test. *Elementary School Journal*, *68*, 259-264. doi: 10.1086/460443

Edwards, M. C., Wirth, R. J., Houts, C. R., & Xi, N. (2012). Categorical data in the structural equation modeling framework. In R. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 195-208). New York, NY: Guilford Press.

Escudero, J. R., & Vallejo, G. (1999) Cuestionario para evaluar las actitudes de los estudiantes de ESO hacia las Matemáticas. *Aula Abierta*, *74*, 193-208.

Estrada, A., & Díez-Palomar, J. (2011). Las actitudes hacia las Matemáticas. Análisis descriptivo de un estudio de caso exploratorio centrado en la Educación Matemática de familiares. *Revista de Investigación en Educación, 9*(2), 116-132.

Fennema, E., & Sherman, J. A. (1976). Fennema-Sherman Mathematics Attitudes Scales: Instruments designed to measure attitudes toward the learning of mathematics by females and males. *Journal for Research in Mathematics Education, 7*(5), 324-326. doi: 10.2307/748467

Gairín, J. (1990). *Las actitudes en educación. Un estudio sobre la educación matemática*. Barcelona: Boixareu Universitaria.

Gil, N., Blanco, L., & Guerrero, E. (2005). El dominio afectivo en el aprendizaje de las Matemáticas. Una revisión de sus descriptores básicos. *UNION. Revista Iberoamericana de Educación Matemática, 2*, 15-32.

González-Pienda, J. A., Fernández-Cueli, M., García, T., Suárez, N., Fernández, E., Tuero-Herrero, E., & Helena da Silva, E. (2012). Diferencias de género en actitudes hacia las matemáticas en la enseñanza obligatoria. *Revista Iberoamericana de Psicología y Salud, 3*(1), 55-73.

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis* (7th. ed.). Upper Sadle River, NJ: Prentice-Hall.

Hambleton, B. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of Item Response Theory*. Newbury Park, CA: Sage.

Hayton, J. C., Allen, D. G., & Scarpello, V. (2004). Factor retention decisions in Exploratory Factor Analysis: A tutorial on Parallel Analysis. *Organizational Research Methods, 7*(2), 191-205. doi: 10.1177/1094428104263675

Hidalgo, S., Maroto, A., & Palacios, A. (2005). El perfil emocional matemático como predictor de rechazo escolar: relación con las destrezas y los conocimientos desde una perspectiva evolutiva. *Revista de Educación Matemática*, *17*(2), 89-116.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1-55. doi: 10.1080/10705519909540118

Kadijevich, D. J. (2008). TIMSS 2003: Relating dimensions of mathematics attitude to mathematics achievement. *Zbornik instituta za pedagoška istraživanja*, *40*(2), 327-346. doi: 10.2298/ZIPI0802327K

McDonald, R. P. (1999). *Test theory: A unified treatment*. Mahwah, NJ: Lawrence Erlbaum Associates.

McLeod, D. B. (1988). Affective issues in mathematical problem solving: Some theoretical considerations. *Journal for Research in Mathematics Education*, *19*, 134-141. doi: 10.2307/749407

McLeod, D. B. (1992). Research on affect in mathematics education: a reconceptualization. En D. Grows (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 575-596). New York: McMillan Publishing Company.

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist, 50*(9), 741-749. doi: 10.1037/0003-066X.50.9.741

Miñano, P., & Castejón, J. L. (2011). Variables cognitivas y motivacionales en el rendimiento académico en Lengua y Matemáticas: un modelo estructural. *Revista de Psicodidáctica, 16*(2), 203-230. doi: 10.1387/RevPsicodidact.930

Miranda, A. (2012). Funcionamiento ejecutivo y motivación en tareas de cálculo y solución de problemas de niños con TDAH. *Revista de Psicodidáctica, 17*(1), 51-72. doi: 10.1387/RevPsicodidact.1839

Morin, A. J. S., Marsh, H. W., & Nagengast, B. (2013). Exploratory Structural Equation Modeling. En G. R. Hancock, y R. O. Mueller (Eds.), *Structural equation modeling: A second course* (2nd ed., pp. 89-102). Charlotte, NC: Information Age Publishing, Inc.

Muñoz, J. M., & Mato, M. D. (2008). Análisis de las actitudes respecto a las matemáticas en alumnos de la ESO. *Revista de Investigación Educativa*, *26*(1), 209-226.

Muthén, B., & Muthén, L. (2013). *MPlus, v. 7.1 [Programa informático]*. Los Angeles, CA: Muthén & Muthén.

Pajares, F., & Miller, D. (1994). Role of self-efficacy and self-concept beliefs in mathematical problem solving: A path analysis. *Journal of Educational Psychology, 86*, 193-203. doi: 10.1037/0022-0663.86.2.193

Perry, C. A. (2011). Motivation and attitude of preservice elementary teachers toward mathematics. Morehead State University. *School Science and Mathematics*, *111*(1), 2-10. doi: 10.1111/j.1949-8594.2010.00054.x

Pett, M. A., Lackey, N. R., & Sullivan, J. J. (2003). *Making sense of factor analysis: The use of factor analysis for instrument development in health care research*. Thousand Oaks, CA: Sage.

Pietsch, J., Walker, R., & Chapman, E. (2003). The relationship among self-concept, selfefficacy, and performance in mathematics during secondary school. *Journal of Educational Psychology, 95*(3), 589-603. doi: 10.1037/0022-0663.95.3.589

Podsakoff, P. M., MacKenzie, S. B., Lee, J., & Podsakoff, N. P. (2003). Common method item bias: A critical review of the literature and recommended remedies. *Journal of Applied Psychology, 88*(5), 879-903. doi: 10.1037/0021-9010.88.5.879

Quiles, M. N. (1993). Actitudes matemáticas y rendimiento escolar. *Comunicación, Lenguaje y Educación, 18*, 115-125.

Reise, S. P., Morizot, J., & Hays, R. D. (2007). The role of the bifactor model in resolving dimensionality issues in health outcomes measures. *Quality of Life Research, 16*(1), 19-31. doi: 10.1007/s11136-007-9183-7

Richardson, F. C., & Suinn, R. M. (1972). The Mathematics Anxiety Rating Scale: Psychometric data. *Journal of Counseling Psychology, 19,* 551-554. doi: 10.1037/h0033456

Sakiz, G., Pape, S. J., & Hoy, A. W. (2012). Does perceived teacher affective support matter for middle school students in mathematics class rooms? *Journal of School Psychology*, *50*, 235-255. doi: 10.1016/j.jsp.2011.10.005

Samejima, F. (1969). Estimation of a latent ability using a response pattern of graded scores. *Psychometrika Monographs, 34* (Suppl. 4). Richmond, VA: Psychometric Society. Retreived from http://www.psychometrika.org/journal/online/MN17.pdf

Samejima, F. (2010). The General Graded Response Model. In M. L. Nering, y R. Ostini (Eds.), *Handbook of polytomous item response theory models* (pp. 43-76). New York: Routledge.

Sandman, R. S. (1980). The Mathematics Attitude Inventory: Instrument and user's manual. *Journal for Research in Mathematics Education, 11*(2), 148-149. doi: 10.2307/748906

Sijtsma, K. (2009). On the use, misuse, and the very limited usefulness of Cronbach's Alpha. *Psychometrika, 74*(1), 107-120. doi: 10.1007/s11336-008-9101-0

Tahara, N. F., Ismailb, Z., Zamanic, N. D., & Adnand, N. (2010). Students' attitude toward mathematics: The use of factor analysis in determining the criteria. *Procedia Social and Behavioral Sciences*, *8*, 476-481. doi: 10.1016/j.sbspro.2010.12.065

Tapia, M., & Marsh, G. E. (2004). An instrument to measure mathematics attitudes. *Academic Exchange Quarterly, 8*(2). Retreived from http://www.rapidintellect.com/AEQweb/cho25344l.htm

Ten Berge, J. M. F., Snijders, T. A. B., & Zegers, F. E. (1981). Computational aspects of the greatest lower bound to reliability and constrained minimum trace factor analysis. *Psychometrika, 46*, 201-213. doi: 10.1007/BF02293900

Timmerman, M. E., & Lorenzo-Seva, U. (2011). Dimensionality assessment of ordered polytomous items with parallel analysis. *Psychological Methods, 16*, 209-220. doi: 10.1037/a0023353

Zumbo, B. D., Gadermann, A. M., & Zeisser, C. (2007). Ordinal versions of coefficients alpha and theta for Likert rating scales. *Journal of Modern Applied Statistical Methods, 6*, 21-29.

Andrés Palacios Picos is Professor at the Department of Psychology (University of Valladolid, Spain), and director of the School of Education of Segovia. He participates in several research groups in R & D, as well as national and regional Educational Innovation Projects. His scientific contributions are numerous, working in priority areas related to didactics of mathematics as well as quantitative analysis in social sciences.

Víctor Arias González is Assistant Professor at the Faculty of Psychology, University of Talca, Chile. His research interests are mainly the study of adaptive behavior and attitudes. He has published several works using the Item Response Theory (concretely, Rasch Model and its variations).

Benito Arias Martínez is Professor at the Department of Psychology, University of Valladolid, Spain. His research interests are centered on calibration of assessment instruments under the postulates of Item Response Theory and Structural Equation Modeling. He has published several papers about adaptive behavior, social anxiety, attitudes, and quality of life.

Note: The matrixes are available upon request from the first author.

Received date: 25-04-2013 Review date: 03-07-2013 Accepted date: 04-09-2013