THE TEACHING OF MATHEMATICS

2006, Vol. IX, 2, pp. 41–51

DEVELOPING TRUSTWORTHY TIMSS BACKGROUND MEASURES: A CASE STUDY ON MATHEMATICS ATTITUDE

Djordje Kadijevich

Abstract. This study, which used a sample of 197,707 students from 46 countries that participated in the TIMSS 2003 project in eight grade, examined whether, for a large number of the TIMSS countries, trustworthy TIMSS measures of several dimensions of mathematics attitude can be developed. By focusing on self-confidence in learning mathematics, usefulness of mathematics, and liking mathematics, it was found that both factor validity and reliability of the measures of these three dimensions derived from the raw data was only attained for the students from the United States. However, when scores concerning the utilized attitudinal statements of all subjects were transformed into Guttman's image form scores, the factor validity and reliability of the three measures utilizing such transformed data was attained for thirty-three countries ($N = 137{,}346$). It was found that for all these thirty-three countries mathematics attitude was mostly saturated by either usefulness of mathematics or self-confidence in learning mathematics. A higher mathematics achievement was found for countries where mathematics attitude was mostly saturated by self-confidence in learning mathematics.

ZDM Subject Classification: C13, C23; *AMS Subject Classification*: 00A35.

Key words and phrases: TIMSS, mathematics attitude, factor validity, reliability, Guttman's image theory.

Mathematics achievement and mathematics attitude are positively related (e.g. [6, 15]). Mathematics attitude is a multidimensional construct whose dimensions can be self-confidence, value of mathematics, enjoyment of mathematics, and motivation [20].

An adequate instrument measuring mathematics attitude should primarily sample cognitive, affective and behavioral domains (taken from [9]). Although the TIMSS 2003 Grade 8 Student Questionnaire did not explicitly and extensively sample these three domains (nor did the project explicitly attempt to assess mathematics attitude and some of its dimensions), the utilized TIMSS statements allow researcher to assess three dimensions of mathematics attitude, namely: self-confidence in learning mathematics (e.g. "I usually do well in mathematics"), usefulness of mathematics (e.g. "I need mathematics to learn other school subjects") and liking mathematics (e.g. "I would like to take more mathematics in school"). These three dimensions can be defined as follows:

(1) self confidence denotes perceived ease, or difficulty, of learning mathematics;

(2) liking mathematics stands for student's affective, emotional and behavioral reactions concerning liking, or disliking, mathematics;

(3) usefulness of mathematics denotes student's beliefs concerting the contribution of mathematics to his/her educational and vocational performance.

Although these three definitions are influenced by the available TIMSS data, they are still given in a general rather than particular context. Recall that some forty years ago Neale viewed student's mathematics attitude in terms of his/her belief that he/she is good or bad at mathematics, his/her liking or disliking of mathematics, his/her belief that mathematics is useful of useless, and his/her tendency to participate in or avoid mathematical activities [17]. As Ma and Kishor remark, mathematics attitude often also includes student's affective responses to the previous two issues concerning perceived ability and usefulness [11].

Apart from the official TIMSS reports analyzing data for all participating countries, just few secondary analyses of the TIMSS data (e.g. [3, 19]) refer to all or most project participants. Furthermore, these official reports do not report the reliabilities of the applied background measures and such a practice has been followed by almost all other TIMSS reports (cf. [8]). Being aware of this inappropriate research practice regarding the scope and trustworthiness of secondary TIMSS analyses, this study examined whether, for a large number of the TIMSS countries, trustworthy TIMSS measures of the three above-mentioned dimensions can be developed from the students' responses to the utilized attitudinal statements concerning mathematics.

Method

Sample

This study used a sample of 197,707 students from forty-six countries that participated in the TIMSS 2003 project in eight grade. Table 1 presents basic facts about the sample by country, where shaded rows point to countries that, because of inappropriate factor structure or indistinguishable item-factor correlations, had to be excluded from the applied analyses at the first (light grey) and second (dark grey) occasions. Note that all students with missing or incomplete data on the examined variables were excluded from this study.

Design

This study utilized a correlative design. The basic variables were: Indicators of Mathematics Attitude (Indicators). The derived variables were: Self-Confidence in Learning Mathematics, Usefulness of Mathematics, and Liking Mathematics.

Instruments

Statements 8a–8d, 8f, 8g, 9a–9e of the TIMSS 2003 Grade 8 Student Questionnaire (see `http://timss.bc.edu/timss2003i/PDF/T03_Student_8.pdf`) were used as Indicators. Item 8e was not used because of its inappropriate loading on the first underlying factor concerning all twelve statements.

Self-Confidence in Learning Mathematics (SCLM) was measured by a 4-item Likert scale administered by means of statements "I usually do well in mathematics", "Mathematics is more difficult for me than for many of my classmates",

Table 1. *Sample size and percentage of students originally assessed by country*

Countries	N	Percentage of students originally assessed
Armenia	4,656	81.3
Australia	4,429	92.4
Bahrain	3,809	90.7
Belgium (Flemish)	4,700	94.6
Botswana	4,386	85.2
Bulgaria	3,618	87.9
Chile	6,130	96.1
Chinese Taipei	5,243	97.5
Cyprus	3,643	91.0
Egypt	5,616	79.2
England	2,581	91.2
Estonia	3,809	94.3
Ghana	3,959	77.6
Hong Kong SAR	4,843	97.4
Hungary	3,131	94.8
Indonesia	5,180	89.9
Iran, Islamic Rep. of	4,533	91.7
Israel	3,858	89.3
Italy	4,054	94.8
Japan	4,627	95.3
Jordan	3,808	84.8
Korea, Rep. of	5,179	97.6
Latvia	3,474	95.7
Lebanon	3,404	89.3
Lithuania	4,187	84.3
Macedonia, Rep. of	3,233	83.0
Malaysia	5,122	96.4
Moldova, Rep. of	3,694	91.6
Morocco	2,160	73.4
Netherlands	2,743	89.5
New Zealand	3,484	91.7
Norway	3,740	90.5
Palestinian Nat'l Auth.	4,735	88.4
Philippines	6,424	92.9
Romania	3,584	87.3
Russian Federation	4,417	94.6
Saudi Arabia	3,477	81.0
Scotland	3,318	94.4
Serbia	3,909	91.0
Singapore	5,922	98.4
Slovak Republic	4,001	94.9
Slovenia	3,291	92.0
South Africa	7,215	80.6
Sweden	3,819	89.7
Tunisia	4,138	83.9
United States	8,424	94.5

"Mathematics is not one of my strengths", and "I learn things quickly in mathematics" (see statements 8a, 8c, 8f and 8g of the Questionnaire; to achieve positive meaning, scoring 1–4 was reversed for items 8a and 8g).

Usefulness of Mathematics (UM) was measured by a 4-item Likert scale administered by means of statements "I think learning mathematics will help me in my daily life", "I need mathematics to learn other school subjects", "I need to do well in mathematics to get into the faculty/university of my choice", "I need to do well in mathematics to get the job I want" (see statements 9a, 9b, 9c and 9e of the Questionnaire; to achieve positive meaning, scoring 1–4 was reversed for all these items).

Liking Mathematics (LM) was measured by a 3-item Likert scale administered by means of statements "I would like to take more mathematics in school", "I enjoy learning mathematics", and "I would like a job that involved using mathematics" (see statements 8a, 8d and 9d of the Questionnaire; to achieve positive meaning, scoring 1–4 was reversed for all these items).

Relevant variables and instruments used in the TIMSS 2003 official study

On the basis the same eleven indicators, the TIMSS 2003 study measured students' self-confidence in learning mathematics and students' valuing mathematics, and represented them by two indices: BSDMSCL–Index of Students' Self-Confidence in Learning Mathematics, and BSDMSV–Index of Students' Valuing Mathematics [10]. The value of BSDMSCL (1, 2 or 3) was derived from the average of student's responses to statements 8a, 8c, 8f and 8g (of the TIMSS 2003 questionnaire mentioned above), by collapsing the calculated averages into three categories. The value of BSDMSV (again 1, 2 or 3) was derived from the average of student's responses to statements 8b, 8d, 9a, 9b, 9c, 9d and 9e (of the questionnaire), also by collapsing the calculated averages into three categories. More detail of the applied procedure can be found in Arora and Ramírez [2] and Mullis et al. [14; pp. 154, 158]. Although Arora and Ramrez are concerned with the reliability of the initial, uncollapsed measure, they do not estimate the reliability of the derived, collapsed measure. Surprisingly, no data on the reliability in question can be found in [12] despite the standard for the IEA studies (including the TIMSS studies) that requires "where appropriate, evidence of the reliability of all scales should be provided in international reports" with a guideline that "scales with low reliability (for example, below 0.7) should be annotated in reports and interpreted with caution" [13, pp. 72, 73].

Statistical analysis and data transformation

For the all examined subjects as well for the subjects from each country, the factor validity and reliability of the three applied measures was determined by the SPSS software that processed weighted data of the whole sample or of the particular country.[1] Because large data sets were analyzed, a principal components factor analysis with *promax* rotation (with Kaiser normalization and kappa=4) was utilized. It was assumed that factor validity was fulfilled when a 3-component solution was applicable, provided that, for close item leadings, the correlation of an item with the desired factor was larger than that with an undesirable factor at a 0.01 level.

To achieve a more precise measurement of Indicators and the three variables (SCLM, VM, and LM), the scores concerning 11 attitudinal statements of all subjects were transformed into Guttman's [5] image form scores[2] and the value of each of the three variables was then represented by the average of the corresponding transformed scores. The transformation applied to raw, initial data was performed by an SPSS macro given in [8].

The reasons for using the three-component solution and applying the transformation in question are summarized below:

- A two-component solution with four self-concept indicators highly loading on one component and the other seven indicators mostly loading on the other that could be applied to the raw data for all forty-six countries was not applicable to each country when its raw data were examined separately.

[1] The TIMSS 2003 international database and its user guide [12] were downloaded from the Internet (`http://timss.bc.edu/timss2003i/userguide.html`). Student's weight was equal to $n *$ *totwgt* / TOTWGT, where n was the sample size, *totwgt* student's total weight given in the official data files, and TOTWGT the sum of all students' individual weights.

[2] This transformation, which eliminates measurement error, is defined by $T = Z(I - R^{-1}U^2)$, where T, Z, I, R and U^2 are, respectively, the following matrices: the matrix of the true results, the matrix of the standardized (and perhaps normalized) initial data, the identity matrix, the matrix of the intercorrelation among the measured variables, and the matrix of the estimate of the variance of measurement error given by $(\operatorname{diag} R^{-1})^{-1}$.

- A three-component solution underlying the three attitudinal dimensions (SCLM, UM and LM) that could be applied to the raw data for all forty-six countries was not applicable to each country when its raw data were examined separately.
- Compared to the EV > 1 (i.e. eigenvalue greater than 1) rule, the scree plot test less overextracts factors [7].
- Components whose eigenvalues are less than 1 (the average variance of all extracted components) may also be included in applied models if "elbows" in scree plots support doing that (see [4]).
- Despite the fact that the third eigenvalue was less than 1 for almost all examined data sets, the three-component model suggested by the scree plot passed the validity test not only for all examined countries (for both the raw and gutmanized data) but also for the majority of these countries when their gutmanized data were examined separately.

Note that the so-called parallel analysis (i.e. parallel principal components factor analysis) proposed by Thompson and Daniel [21], comparing the gutmanized raw data vs. the gutmanized random data for all thirty-three countries, yielded the following eigenvalues: **8.296, 1.347, 0.421**, 0.237, 0.214, 0.165, 0.128, 0.087, 0.075, 0.028, 0.000 versus **9.990, 0.130, 0.125**, 0.114, 0.105, 0.102, 0.095, 0.091, 0.088, 0.083, 0.075, respectively.

Results

Table 2 reports the factor validity and reliability of the three derived variables for initial, raw data. Having in mind the usually assumed reliability cut-off of 0.70 (e.g. [18]), the values of the three variables may only be confidently used for the students from the United States. More detail on this unfavorable outcome can be found in Appendix I available on the Internet (see `www.mi.sanu.ac.yu/~djkadij/AppendixI.pdf`) where full results for the whole 46-country sample and a 10-country random sample are given. Note that Cronbach's alpha reliability of Mathematics Attitude, represented by the average of the eleven utilized scores, was less than 0.70 only in Philippines (0.69), Ghana (0.69) and Botswana (0.67).

The applied data transformation considerably improved the reliability in question because, for each of the three measures, Cronbach's alpha was not below 0.89 and was usually around 0.95. Furthermore, the transformation totally improved the factor validity of the three measures because this validity could be attained for thirty-three countries[3] listed in the un-shaded rows of Table 1.

[3] These countries were found by means of three analyses. Having transformed raw data of all students from 46 countries, the first analysis evidenced that factor validity was not attained for 10 countries, and these countries were excluded from further analysis. Having transformed raw data of all students for the remaining 36 countries, the second analysis evidenced that factor validity was not attained for 3 countries, and these countries were also excluded from further analysis. Finally, having transformed raw data of all students from remaining 33 countries, the third analysis evidenced that factor validity was attained for all these countries.

Table 2. *Factor validity and reliability of the applied measures by country*

Country	Factor validity	Chronbach's alpha reliability of		
		SCLM	UM	LM
Armenia	Unattained	0.55	0.68	0.72
Australia	Unattained	0.83	0.77	0.78
Bahrain	Unattained	0.72	0.68	0.77
Belgium (Flemish)	Unattained	0.81	0.70	0.81
Botswana	Unattained	0.45	0.61	0.59
Bulgaria	Unattained	0.71	0.67	0.76
Chile	Unattained	0.64	0.64	0.72
Chinese Taipei	Unattained	0.82	0.77	0.79
Cyprus	Unattained	0.77	0.69	0.76
Egypt	Unattained	0.52	0.55	0.52
England	Unattained	0.80	0.70	0.72
Estonia	Unattained	0.85	0.74	0.71
Ghana	Unattained	0.45	0.62	0.52
Hong Kong SAR	Unattained	0.84	0.76	0.77
Hungary	Unattained	0.84	0.62	0.73
Indonesia	Unattained	0.53	0.64	0.62
Iran, Islamic Rep. of	Unattained	0.45	0.68	0.71
Israel	Unattained	0.74	0.69	0.72
Italy	Unattained	0.85	0.69	0.79
Japan	Unattained	0.78	0.66	0.71
Jordan	Unattained	0.63	0.69	0.72
Korea, Rep. of	Unattained	0.86	0.71	0.74
Latvia	Unattained	0.79	0.68	0.67
Lebanon	Unattained	0.51	0.62	0.65
Lithuania	**Attained**	0.79	0.69	0.68
Macedonia, Rep. of	Unattained	0.68	0.73	0.79
Malaysia	Unattained	0.68	0.67	0.66
Moldova, Rep. of	Unattained	0.65	0.74	0.69
Morocco	Unattained	0.66	0.65	0.69
Netherlands	Unattained	0.85	0.64	0.66
New Zealand	Unattained	0.80	0.74	0.73
Norway	Unattained	0.81	0.76	0.75
Palestinian Nat'l Auth.	Unattained	0.60	0.68	0.68
Philippines	Unattained	0.32	0.67	0.60
Romania	Unattained	0.61	0.67	0.74
Russian Federation	**Attained**	0.83	0.69	0.74
Saudi Arabia	Unattained	0.59	0.70	0.76
Scotland	Unattained	0.79	0.72	0.71
Serbia	**Attained**	0.80	0.66	0.72
Singapore	Unattained	0.82	0.68	0.83
Slovak Republic	Unattained	0.80	0.66	0.71
Slovenia	**Attained**	0.77	0.69	0.76
South Africa	Unattained	0.43	0.72	0.68
Sweden	Unattained	0.83	0.70	0.72
Tunisia	Unattained	0.71	0.70	0.73
United States	**Attained**	0.83	0.71	0.79

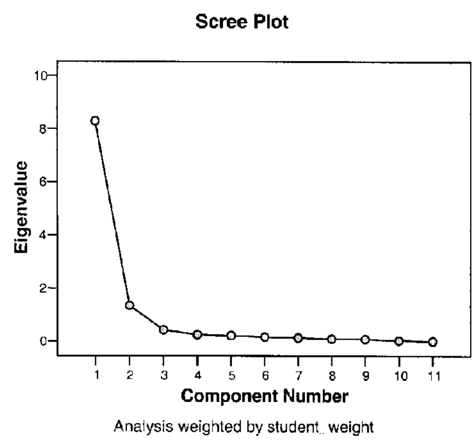
The results of the factor analysis applied to the transformed data for all countries and Serbia are presented in Tables 3 and 4, respectively. These results evidence that mathematics attitude is mostly saturated by usefulness of mathematics for the whole sample of thirty-three countries, whereas this attitude is primarily saturated by self-confidence in learning mathematics for Serbia. The results for a 10-country random sample given in Appendix II (available at the Internet at `www.mi.sanu.ac.yu/~djkadij/AppendixII.pdf`) evidence the following: mathematics attitude is mostly saturated by usefulness of mathematics for Bahrain, Cyprus, Israel and Romania, whereas this attitude is primarily saturated by self-confidence in learning mathematics for Belgium (Flemish), Chinese Taipei, Hungary, Malaysia, Norway and Russian Federation. This pattern that mathematics attitude is mostly saturated by either usefulness of mathematics or self-confidence in learning mathematics applies to all examined countries.

Discussion

Three important findings emerged from this study. First, the factor validity and reliability of the three applied measures based upon the raw data was only attained for the students from the United States. Second, the factor validity and reliability of these measures utilizing the transformed data was attained for thirty-three countries. Third, mathematics attitude was mostly saturated by either usefulness of mathematics or self-confidence in learning mathematics.

Table 3. *All countries: Factor pattern matrix and eigenvalues for three components*

Item	Component		
	1	2	3
I need mathematics to learn other school subjects	0.96		
I need to do well in mathematics to get into the faculty/university of my choice	0.93		
I need to do well in mathematics to get the job I want	0.91		
I think learning mathematics will help me in my daily life	0.80		
Mathematics is more difficult for me than for many of my classmates		1.06	
Mathematics is not one of my strengths		.96	
I learn things quickly in mathematics		.82	
I usually do well in mathematics		.80	
I would like to take more mathematics in school			0.88
I would like a job that involved using mathematics			0.71
I enjoy learning mathematics	0.30		0.70
Eigenvalue	8.30	1.35	0.42
% of Variance	75.4	12.2	3.8

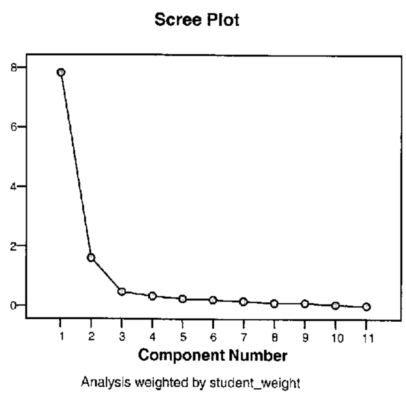


The appropriateness of the three components was indicated by the "elbow" on the scree plot. Only matrix elements with absolute values greater than 0.30 are displayed. There were 9 (16%) non-redundant residuals with absolute values greater than 0.05 for the two-component solution, whereas there were just 5 (9%) such residuals for the three-component solution.

The first finding requires researcher to primarily make use of the construct of mathematics attitude represented by the average of the eleven raw scores. Of course, the individual measures of the three variables (SCLM, UM and LM) can be used for many countries but not as the dimensions of mathematics attitude. It is important to realize that the reliability of officially proposed index BSDMSCL (Index of Students' Self-Confidence in Learning Mathematics) is problematic because, as Table 2 evidences, the reliability of its initial, uncollapsed measure SCLM is below 0.70 for many countries.

Table 4. *Serbia: Factor pattern matrix and eigenvalues for three components*

Item	Component		
	1	2	3
Mathematics is more difficult for me than for many of my classmates	1.06		
Mathematics is not one of my strengths	0.95		
I learn things quickly in mathematics	0.85		
I usually do well in mathematics	0.83		
I need to do well in mathematics to get the job I want		0.96	
I need mathematics to learn other school subjects		0.95	
I need to do well in mathematics to get into the faculty/university of my choice		0.83	
I think learning mathematics will help me in my daily life		0.82	
I would like to take more mathematics in school			0.83
I would like a job that involved using mathematics			0.82
I enjoy learning mathematics	0.32		0.72
Eigenvalue	7.83	1.61	0.47
% of Variance	71.2	14.7	4.2



The appropriateness of the three components was indicated by the "elbow" on the scree plot. Only matrix elements with absolute values greater than 0.30 are displayed. There were 13 (23%) non-redundant residuals with absolute values greater than 0.05 for the two-component solution, whereas there were just 6 (10%) such residuals for the three-component solution.

The second finding, which reestablishes a high research value of the applied Guttman's transformation [5], paves the way for developing trustworthy TIMSS measures. Apart from [8], this transformation has, to the author's knowledge, never been applied in the TIMSS context. However, bearing in mind that the TIMSS International Study Center used the eleven raw scores to create two indices—BSDMSCL obtained from SCLM, whereas BSDMSV, Index of Students' Valuing Mathematics, derived from $(4\text{UM}+3\text{LM})/7$—one may question the applied three-factor solution. This is because, if a two-factor solution was more appropriate, it introduced more complexity than needed, probably unnecessarily calling for the transformation in question. Let us repeat that the two-component solution with four self-concept indicators highly loading on one factor and the other seven indicators mostly loading on the other was not applicable to each country when its raw data were examined separately. A random sample of ten[4] out of 33 countries clearly evidenced this fact because, apart from the raw data for Israel (yet with three eigenvalues greater than 1), neither raw, nor transformed data supported such a two-component solution (see Appendix II at the Internet).

The third finding that mathematics attitude was mostly saturated by either usefulness of mathematics or self-confidence in learning mathematics initiated an additional analysis involving mathematics achievement (represented by the mean of the five plausible values given in the official TIMSS data; the weighted data were again used). This analysis showed the following:

- Mathematics attitude was mostly saturated by usefulness of mathematics for eleven countries: Australia, Bahrain, Bulgaria, Chile, Cyprus, Indonesia, Israel, Rep. of Macedonia, Rep. of Moldova, Romania, and Scotland. Mathematics achievement in these countries ranged from 388 for Chile to 508 Australia, with the mean of 459 and standard deviation of 41 (countries were treated equally).
- Mathematics attitude was mostly saturated by self-confidence in learning mathematics for the remaining twenty-two countries: Belgium (Flemish), Chinese Taipei, England, Estonia, Hong Kong SAR, Hungary, Italy, Japan, Jordan, Rep. of Korea, Latvia, Lithuania, Malaysia, Morocco, New Zealand, Norway, Russian Federation, Serbia, Slovak Republic, Sweden, Tunisia, and United States. Mathematics achievement in these countries ranged from 392 for Morocco to 591 for Rep. of Korea, with the mean of 507 and standard deviation of 52 (countries were again treated equally).
- Differences in mathematics achievement between the two groups of countries (459 vs. 507) were significant at a 0.05 level ($t = 2.68$, $df = 31$, $p = 0.012$).

Although the major saturation in question can not be a definitive sign of such and such mathematics education, it may be said that the societal context of mathematics education tend to primarily challenge either the utility of mathematics or the mathematical competency of their students, and that challenging the latter would result in better mathematics achievement.

The reader may ask whether the factor analysis is the best possible tool for establishing the validity of the construct in question and its dimensions. Although a thoughtful validation of an instrument requires researcher to establish, among other issues, its convergent and discriminant validity involving confirmatory factor analyses [1], research studies of this sort usually make only use of a factor or cluster analysis to tests whether conceptually postulated dimensions are reflected in empirically derived factors or clusters (see [16]). The fact that even this requirement is not easy to achieve is evidenced by the examined TIMSS raw data, which, without the applied transformation, can not be confidently used. (Recall that the official TIMSS 2003 two-component solution with four self-concept indicators high-

[4] These ten countries were Bahrain, Belgium (Flemish), Chinese Taipei, Cyprus, Hungary, Israel, Malaysia, Norway, Romania, and Russian Federation.

ly loading on one factor and the other seven indicators mostly loading on the other was not applicable to each country when its raw data were examined separately.) The guttmanization of the data however triggers another important question "Does such a data fudging preserve the validity of the outcomes?" Because, as already explained in footnote 2, the applied transformation eliminated the measurement error regarding the eleven attitudinal indicators, the measurement did become more reliable. However, a high reliability does not consequentially imply a high validity: despite their reliability, the applied measures may be biased not representing the construct that are intended. There were no gold standard measures in this research against which the reliability and validity of the construct in question and its three dimensions were to be examined [10]. Despite that, the questioned validity can still be established not only because of the third finding that mathematics attitude was mostly saturated by either self-confidence in learning mathematics or usefulness of mathematics, but also because of the additional finding that countries where mathematics attitude was mostly saturated by self-confidence in learning mathematics had, on the average, higher mathematics achievement than countries where mathematics attitude was primarily saturated by usefulness of mathematics. It is the guttmanized data that made possible the discovery of this important finding.[5]

To summarize: This study evidenced that the raw TIMSS 2003 international data might be of little use concerning the confident measurement of some dimensions of mathematics attitude. The things considerably improved when the raw scores were transformed into Guttman's image form scores. Because of that, researchers should develop and use, whenever needed and possible, their own trustworthy measures of TIMSS background variables. By doing that, important, previously uncovered patterns may emerge such as "mathematics attitude is primarily saturated by either usefulness of mathematics or self-confidence in learning mathematics", which may help us explain differences in mathematics achievement across countries. Further research concerning the three dimensions of mathematics attitude may search for one of them that is mostly related to mathematics achievement.

Acknowledgement. This contribution resulted from the author's work on the project 144050A "History and philosophy of mathematics and natural sciences" funded by the Serbian Ministry of Science and Environment Protection.

REFERENCES

1. Ahire, S. L. & Devaraj, S. (2001), *An empirical comparison of statistical construct validation approaches*, IEEE Transactions on Engineering Management, 48, 3, 319–329.
2. Arora, A. & Ramírez, M. J. (2004), *Developing indicators of educational contexts in TIMSS* (paper presented at the 1st IEA International Research Conference, Lefkosia, Cyprus, May 11–13, 2004). Retrieved 13 February, 2007, from www.iea.nl/fileadmin/user_upload/IRC2004/Arora_Ramirez.pdf
3. Elley, W. B. (2005), *Evaluating students' achievements: How TIMSS-R contributed to education in eighteen developing countries*, Prospects, 35, 2, 199–212.

[5]If the guttmanized data are obtained by applying the transformation in question to the raw data of each country separately, the obtained 3-component structure may not apply as in the case of Serbia.

4. Everitt, B. S. (1996), *Making Sense of Statistics in Psychology*, Oxford: Oxford University Press.
5. Guttman, L. (1953), *Image theory for the structure of quantitative variates*, Psychometrika, 21, 3, 277–296.
6. Hembree, R. (1992), *Experiments and relational studies in problem solving: A meta-analysis*, Journal for Research in Mathematics Education, 23, 3, 242-273.
7. Henson, R. K. & Roberts, J. K. (2006), *Use of exploratory factor analysis in published research: Common errors and some comment on improved practice*, Educational and Psychological Measurement, 66, 3, 393–416.
8. Kadijevich, Dj. (2003), *Examining mathematics attitude in a TIMSS 2003 pilot research*, Annals of the Institute for Educational Research, 35, 96–102. Retrieved 13 February, 2007, from `www.mi.sanu.ac.yu/~djkadij/zipi03.pdf`
9. Kay, R. (1993), *An exploration of theoretical and practical foundations for assessing attitudes toward computers: the Computer Attitude Measure (CAM)*, Computers in Human Behavior, 9, 4, 371–386.
10. Lewis, R. J. (1999), *Reliability and validity: meaning and measurement*, Paper presented at the 1999 Annual Meeting of the Society for Academic Emergency Medicine (SAEM) in Boston, Massachusetts. Retrieved 13 February, 2007, from `www.ambpeds.org/ReliabilityandValidity.pdf`
11. Ma, X. & Kishor, N. (1997), *Attitude toward self, social factors, and achievement in mathematics: A meta-analytic review*, Educational Psychology Review, 9, 2, 89–120.
12. Martin, M. O. (Ed.) (2005), *TIMSS 2003 User Guide for the International Database*, Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College. Retrieved 13 February, 2007, from `http://timss.bc.edu/timss2003i/userguide.html`
13. Martin, M. O., Rust, K. & Adams, R. J. (Eds.), with contributions from Caldwell, N., Foy, P., Martin, M. O., Mullis, I. V. S. & Sibberns, H. (1999), *Technical Standards for IEA Studies*, Amsterdam: International Association for the Evaluation of Educational Achievement.
14. Mullis, I. V.S., Martin, M. O., Gonzalez, E. J. & Chrostowski, S. J. (2004), *TIMSS 2003 International Mathematics Report*, Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College. Retrieved 13 February, 2007, from `http://timss.bc.edu/timss2003i/mathD.html`
15. Mullis, I. V. S., Martin, M. O., Gonzalez, E., O'Connor, K. M., Chrostowski, S. J., Gregory, K. D., Garden, R. A. & Smith, T. A. (2001), *TIMSS 1999 Benchmarking Mathematics Report*, Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College. Retrieved 13 February, 2007, from `http://timss.bc.edu/timss1999b/mathbench_report/t99b_math_report.html`
16. Munby, H. (1997), *Issues of validity in science attitude measurement*, Journal of Research in Science Teaching, 34, 4, 337–341.
17. Neale, D. C. (1969), *The role of attitudes in learning mathematics*, Arithmetic Teacher, 16, 8, 631–640.
18. Nunnally, J. C. (1978), *Psychometric Theory* (2nd ed.), New York: McGraw Hill.
19. Shen, C. (2002), *Revisiting the relationship between students' achievement and their self-perceptions: A cross-national analysis based on the TIMSS 1999 data*, Assessment in Education, 9, 2, 161–184.
20. Tapia, M. & Marsh II, G. E. (2005), *An instrument to measure mathematics attitudes*, Academic Exchange Quaterly, 8, 2. Retrieved 13 February, 2007, from `www.rapidintellect.com/AEQweb/cho25344l.htm`
21. Thompson, B. & Daniel, L. G. (1996), *Factor analytic evidence for the construct validity of scores: A historical overview and some guidelines*, Educational and Psychological Measurement, 56, 2, 197–208

Djordje Kadijevich, TIMSS NRC for Serbia in 2001, Megatrend University and Mathematical Institute SANU, Belgrade, Serbia

E-mail: `djkadijevic@megatrend-edu.net`