Eurasia Journal of Mathematics, Science & Technology Education, 2016, **12**(4), 965-974

PISA and TIMSS Science Score, Which Clock is More Accurate to Indicate National Science and Technology Competitiveness?

Wei-Zhao Shi, Xiqin He, Yan Wang, Zeng-Guang Fan & Liangdong Guo
University of Science and Technology Liaoning, CHINA

•*Received 16 September 2015*•*Revised 1 November 2015* •*Accepted 20 November 2015*

In 2015, PISA and TIMSS are coming up to us together. In this study, the data from PISA and TIMSS are used to investigate that which one is a better indicator of national science and technology (S&T) competitiveness? Number of S & T journal articles (per million people) is used as a measure to represent the national S&T competitiveness. Average IQ of the population, research and development (R&D) expenditure (% of GDP) and number of R&D researchers and technicians which affect the national competitiveness in S&T were also investigated. The study shows that PISA science scores would more significantly indicate national S&T competitiveness than TIMSS. Moreover, the study also shows a strong link between competence in S&T and IQ, research and development expenditure (% of GDP) or number of research and development researchers and technicians. Some possible micro-foundations of these relationships are discussed, and policy implication is clear.

Keywords: National S&T competitiveness, PISA, TIMSS

INTRODUCTION

> "A man with a watch knows what time it is. A man with two watches is never sure" (Segal's law)

Around the world, one area of interest in education is comparative studies in educational achievement, in particular, in mathematics, science and reading. There are two well-subscribed programs involving science, namely, PISA and TIMSS. In the same year 2015, PISA (The Programme for International Student Assessment) and TIMSS (Trends in International Mathematics and Science Study) are coming up to us together. PISA is a triennial international survey which aims to evaluate education systems worldwide by testing the skills and knowledge of 15-year-old students since 2000, while TIMSS has been conducted on a regular 4-year cycle since 1995, and measures the trends in mathematics and science achievement at the fourth and eighth grades. Grades are the main indicator of ability and performance that can

Correspondence: Wei-Zhao Shi,
School of Science, University of Science and Technology Liaoning, 114051, Anshan, CHINA
E-mail: vcshih@aliyun.com
doi: 10.12973/eurasia.2016.1239a

ISSN: 1305-8223

have long-term consequences for pupil achievement and thereby on future employment perspectives, High performing school systems prepare their students for these knowledge-based jobs (Tucker, 2011). Able students with a good understanding of science form a pool of future engineers and scientists. It has been taken for granted that assessment achievement is a good measure of quality of human capital (Lee & Barro, 2001), and national competence in science and technology primarily depend on the quality of human capital. The importance of human capital at the country level is supported by the observation of large differences in labor productivity between countries (Hall & Jones, 1999). PISA and TIMSS, which clock is more accurate to indicate national S&T (science & technology) competitiveness?

Qualitative descriptions of S&T competitiveness, when considered, generally come from one of two sources: quality of human capital or R&D inputs. Average IQ of the population provides another measure of quality of human capital (Meisenberg & Lynn, 2011). In this study, the researcher not only considered the education influence but also focused the other factors which affect the national competitiveness in S&T. Whether and to what extent these factors are related to S&T competitiveness? The results are very useful for every country, regardless of this country belongs to developed countries or developing countries. This paper provided a guider to national policy makers.

State of the literature

- TIMSS focused on the extent to which students have mastered as they appear in school curricula, PISA aimed to capture the ability to use knowledge and skills to meet real-life challenges.
- Able students with a good assessment achievement of science form a pool of future engineers and scientists, the average IQ of the population is also a good measure of quality of human capital.
- Qualitative descriptions of S&T competitiveness, generally come from two sources: quality of human capital or R&D inputs

Contribution of this paper to the literature

- The study shows PISA is a better measure of competence in S & T than TIMSS, this because PISA aims to assess scientific literacy not only scientific knowledge.
- IQ is weaker than science score in predicting competence in S&T. The reason is scientific literacy can be obtained through high quality educational system rather than natural IQ.
- R&D input is also an important determinant to the competence in S & T, and a larger proportion of R&D expenditure cost on the researchers' salary payments.

METHODOLOGY

Since the number of scientific and technical journal articles (per million people) which shows the activities involved in research and development can better represent the competence of science and technology (Chen & Luoh, 2010), the researcher used it to investigate the relationships between PISA or TIMSS science scores and competence in science and technology, would like to see whether there is a strong link between them. Figure 1 shows the ranking of 59 countries in number of scientific and technical journal articles per million people. The researcher tested the link between science scores and competence of (S&T) while other variables such as average IQ of the population, research and development researchers (per million people) and research and development expenditure (% of GDP) which are important to competence in S&T were also included in the regression model.

Data

PISA 2009 and TIMSS 2011 science scores are publicly available. PISA results are available from OECD, http:// www.pisa.oecd.org. In PISA 2009, students in 65 countries/regions took part in test. And eighth grade students' science score data from the TIMSS 2011, which is available from U.S. National Center for Education Statistics, http://nces.ed.gov/timss/table11_5.asp. In TIMSS 2011, students in 42 countries/regions took part in test. The logarithmic transformation was used

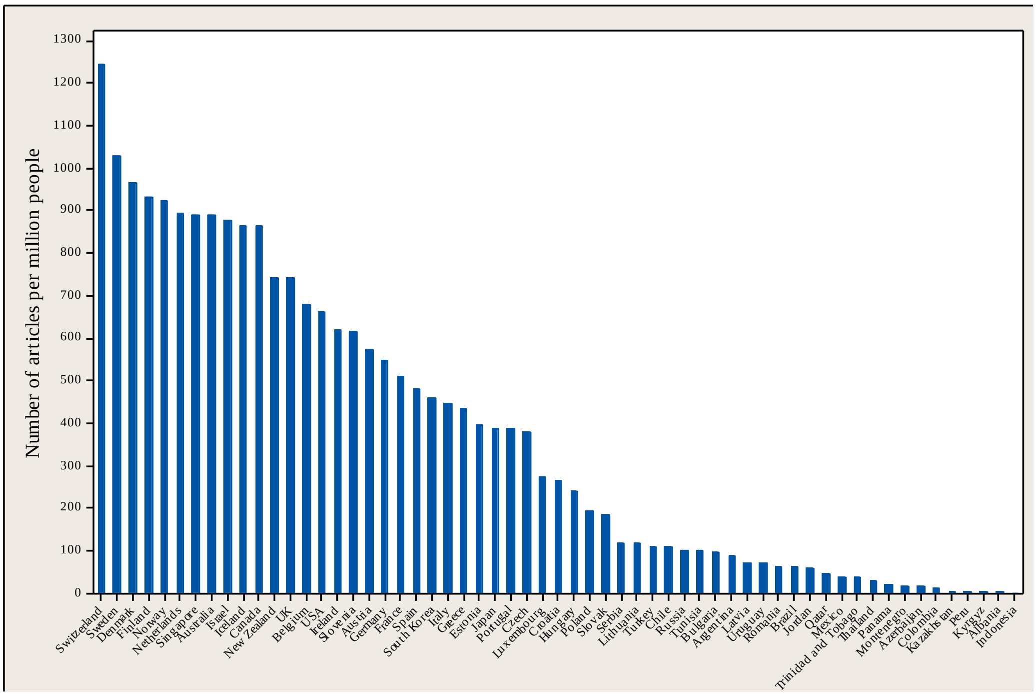

Figure 1. Number of scientific and technical journal articles per million people ranking of 59 countries

because of the highly skewed nature of students' science score worldwide, which approximates to a normal distribution in the logarithmic form.

The national average IQ data come from Meisenberg & Lynn (2011). World population (millions) is available from United Nations statistical databases. The number of Scientific and Technical journal articles, the number of researchers in research and development (R&D) per million people and the R&D expenditure (% of GDP) are obtained from World Development Index constructed by the World Bank. All the data is shown in the appendix.

Analysis of relationship between students' science scores and national competitiveness in S&T

In PISA 2009, about 475,000 students from over 17,000 schools in 65 countries/regions took part in a two-hour test. Because of the unavailability of Hong Kong-China's, Macau-China's, Shanghai-China's, Chinese Taipei (Taiwan)'s, Liechtenstein's and Dubai's article or population data, 59 countries are taken into account. In TIMSS 2011, students in 42 countries/regions took part in test. Because of the unavailability of Hong Kong-China's, Chinese Taipei (Taiwan)'s and Palestinian's article or population data, 39 countries are taken into account.

Figure 2 shows a plot of PISA science scores and S&T journal articles per million people for the 59 countries. The R^2 as a measure of the proportion of variance explained is 0.772, this shows a good evaluate indicator for competence of S&T. The line shown in Figure 2 is the line of best fit for the data points. Figure 3 shows a plot of TIMSS science test scores and S&T journal articles per million people for the 39 countries. The R^2 as a measure of the proportion of variance explained is 0.590. From comparison between Figures 2 and 3, we can conclude that PISA science scores is a better measure of national competitiveness in S&T than TIMSS.

It can be seen from both Fig.2 and Fig.3, generally, Middle East and South Eastern Asia and Latin American countries are fitted nearly perfectly; most of these countries' science scores are below 450 in PISA or 480 in TIMSS. Western countries performed a little better than predicted. Eastern European countries (such as Russian Federation, Estonia, Hungary, Poland, Slovak Republic, Lithuania and

Latvia) and Confucian Asian countries (such as Japan, South Korea and Singapore) generally performed a little poorer than predicted. The results are similar with previous studies (Wu, 2009). This observation prompted an investigation of whether other factors relating to characteristics of Western and Eastern European/Asian countries may have an impact.

Other possible factors impacting S&T competence

PISA shows a better evaluate indicator for national competitiveness in S&T than TIMSS, however, average IQ of the population provides another measure of quality of human capital. It should be considered. And some other factors impacting S&T competence also should be factored in. For instance, Israel's Performance on PISA science was lower, but their research and development expenditure is the highest in the world at 4.27% (data from World Development Index constructed by the World Bank). This may be the reason why Israel has the strong competence of S&T. As the same, Nordic countries' (Iceland, Norway, Denmark, Sweden, and Finland) number

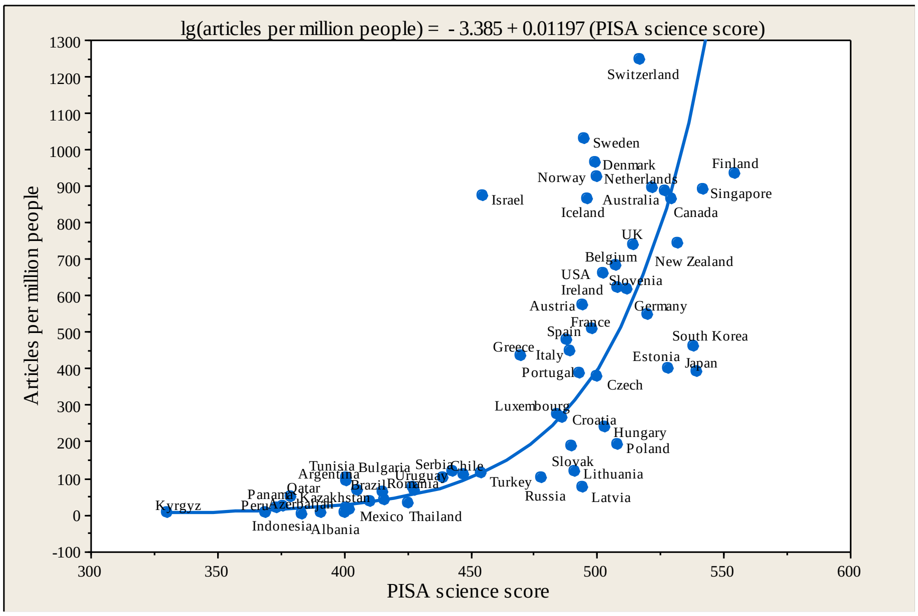

Figure 2. PISA Science score and S&T journal articles per million people

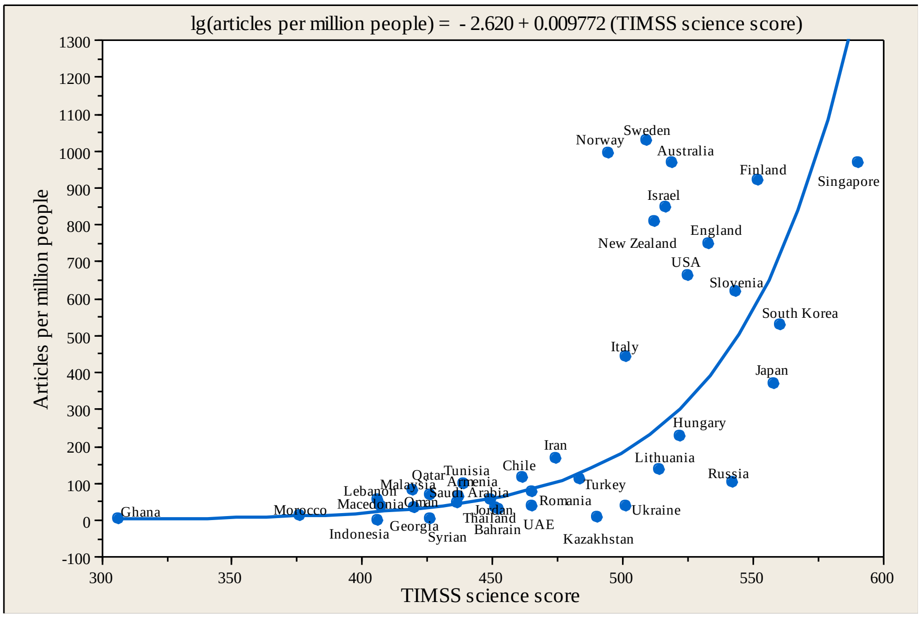

Figure 3. TIMSS Science score and S&T journal articles per million people

of researchers in R&D per million people are the top countries in the world, average IQ of the population and the R&D expenditure (% of GDP) are also much higher in the world. Maybe these factors contribute to the Nordic countries' higher S&T competence. The number of S&T journal articles per million people in Switzerland is the highest in the world. This may be due to the combination of four important factors: higher research and development (R&D) expenditure, more researchers in R&D, higher quality of education, and higher average IQ of the population.

So we must consider these factors which also influence the S&T competence: average IQ of the population, the number of researchers in R&D per million people and R&D expenditure (% of GDP). The researcher would like to see whether there is still a strong link between science scores and S&T competence while adding the other three important factors.

Regression Model

Standard regression model is considered as follows:

$$\lg(artical) = \beta_0 + \beta_1 score + u \quad (1)$$

$$\lg(artical) = \beta_0 + \beta_1 score + \beta_2 reseacher + u \quad (2)$$

$$\lg(artical) = \beta_0 + \beta_1 score + \beta_2 reseacher + \beta_3 IQ + u \quad (3)$$

$$\lg(artical) = \beta_0 + \beta_1 score + \beta_2 reseacher + \beta_3 IQ + \beta_4 \exp enditure + u \quad (4)$$

Where article, score, researcher, IQ, expenditure and u are number of S&T journal articles per million people, PISA science scores, researchers in R&D (per million people), Intelligence tests, research and development expenditure (% of GDP) and the random error term, respectively. Here PISA science scores are used, because it is a better measure of national competitiveness in S&T than TIMSS.

RESULTS

Due to the unavailability of data on the number of researchers in R&D per million people, IQ or the R&D expenditure (% of GDP) in some countries in the year 2009, there are 42 countries remained. Table.1 shows the correlations of four factors and log-transformed articles per million people.

A number of observations can be made:

1. All variable including log-transformed articles per million people, research and development (R&D) expenditure, researchers in R&D, science scores, and average IQ of the population, form a positive manifold. However, the correlations between R&D expenditure (% of GDP) and number of researchers in R&D per million people are higher than any of the other correlations in the table.
2. The science scores is more highly correlated with lg(articles) than the other factors, this suggests the science scores is a most important determinant of the S&T competence.
3. IQ is related more closely to the science scores than other factors. This shows that higher IQ students can get better grades.

Tables 2 shows the results of regression models in which various outcomes are predicted by the four factors: science scores, researchers in R&D, average IQ of the population research and development (R&D) expenditure. The results of Tables 2 confirm that lg(articles) are related more closely to science scores than other three factors.

When the number of researchers in R&D per million people added in the analysis, we can see clearly the correlation between PISA science scores and the number of scientific and technical journal articles (per million people) remains significant (see column (2) of Table 2). It is worth noting that under this model, both factors contribute significantly to explain the number of scientific and technical journal articles (per million people). In column (3) of Table 2, the average IQ of the population is added into the regression model. The significant impact of science scores remains, but the number of researchers in R&D per million people is no longer related to competence. However, if average IQ of the population and the R & D expenditure (% of GDP) are retained as the only two measures of the number of scientific and technical journal articles (per million people), the result in column (5) suggest the two factors are very important variables that contribute to national S &T competence. IQ of the population and R & D expenditure are both significant at 1%.

DISCUSSION AND CONCLUSION

The most important finding in this study is that differential performances of countries in S &T competence can be largely accounted for by two kinds of factors. In particular, quality of human capital is the most significant factor. The other factor is R&D inputs. Average IQ of the population and educational test score provide measures of quality of human capital. R&D inputs are embodied in the number of researchers in R&D per million people and R&D expenditure (% of GDP).

Firstly, the researcher investigated whether students' science score is a good indicator of the competence in S&T. The study shows PISA is a better measure of competence in S & T than TIMSS. What makes different? Both PISA and TIMSS assess students' achievement levels in science. There has been considerable interest in comparing the two surveys, as the results from these two surveys have not always been consistent (Brown et al, 2007). So it is essential to gain a clear understanding of what each survey assesses. Barry McGaw, the Director for Education of the OECD, characterized the difference as TIMSS being interested to discover, "what science have you been taught and how much have you learned?", while for PISA it was "what can you do with the science you have been taught?". Rather than assessing the same subjects as TIMSS, PISA aims to assess scientific literacy. As a mission statement, it is claimed that:

> The prime aim of the OECD/PISA assessment is to determine the extent to which young people have acquired the wider knowledge in reading literacy, mathematical literacy and scientific literacy that they will need in adult life.(OECD, 2004)

Furthermore, average IQ of the population as another measure of quality of human capital is also obvious to indicate competence in S&T. The reason we can find from Table 1, the correlation coefficient (0.805) between educational test score and IQ of the population is quite higher. Studies repeatedly show that performance on intelligence tests is correlated with school achievement (Brody, 1997; Sattler, 2002). At the individual level within countries, correlations between IQ tests and school achievement tests are typically between 0.5 and 0.7 (Jencks et al, 1972; Jensen, 1998; Mackintosh, 2011), but can be as high as 0.8 (Deary et al, 2007). At the country level, correlations between the results of IQ tests and scholastic assessments are in the vicinity of 0.9 (Lynn & Meisenberg, 2010). However, Table1 and Table 2 show that IQ is weaker than science test score in predicting competence in S&T. Why? A reasonable explanation is that scientific literacy can be obtained through high quality educational system rather than natural IQ. We must also note

Table 1. Correlation coefficient

	researcher	**expenditure**	**science scores**	**lg(articles)**
expenditure	0.873			
science scores	0.774	0.722		
lg(articles)	0.769	0.732	0.856	
IQ	0.740	0.681	0.805	0.762

Table 2. Regression models predicting lg (Article)

lg (article)					
science scores	0.011783 [10.48] **	0.008958 [5.27] **	0.007909 [3.89] **	0.007730 [3.76] **	
researcher		0.00007799 [2.15] *	0.00006672 [1.74]	0.00003741 [0.70]	
IQ			0.01491 [0.0847]	0.01456 [0.92]	0.05576 [3.89] **
expenditure				0.07460 [0.78]	0.23115 [3.15] **
Constant	-3.3097 [-6.06] **	-2.1775 [-2.93] **	-3.068 [-2.55] *	-2.978 [-2.46] *	-3.329 [-2.56] *
Observations	42	42	42	42	42
R-squared	0.733	0.761	0.767	0.770	0.666

*Notes: Absolute value of t statistics in parentheses; * significant at 5%, ** significant at 1%.*

that the relationship between IQ scores and school achievement is an imperfect one, with many exceptions to the rule. Although students with high IQs typically perform well in school, we cannot say conclusively that their high achievement is actually the result of their intelligence. Intelligence probably does play an important role in school achievement, but quality of instruction is also involved. Therefore, to a certain extent, the difference between school achievement and IQ can be used as a measure for the quality of the educational system (Meisenberg & Lynn, 2011).

Moreover, the study also shows that R&D input is another important determinant to the competence in S & T. R&D inputs are embodied in the number of researchers in R&D per million people and R&D expenditure (% of GDP). From the study, we can see that they are both important determinants to the competence in S & T. The correlations between R&D expenditure (% of GDP) and number of researchers in R&D per million people are higher than any of the other correlations in the Table 1. One plausible explanation is that a larger proportion of R&D expenditure cost on the researchers. The study of Goolsbee (1998) showed that the majority of R&D spending is actually salary payments for R&D workers and the supply of this scientific and engineering talent is quite inelastic. N.S.F. (1995) documents that between 45 and 83% of total spending is wages and benefits of scientific personnel (depending on how one counts overhead which includes individual benefits). A reasonable approximation for the total share might be 2/3. So in 1995 the US government spent almost $70 billion on R&D, $45 billion of that was wages and benefits for R&D workers.

Higher quality of human capital and more R&D inputs in Confucian Asian countries, however, Confucian Asian countries (such as Japan, Korea and Singapore) generally performed a little poorer than predicted. A possible explanation is the lower self-concept. Report based on OECD (2007) showed that the lowest scoring countries on science self-concept were Confucian Asian countries, especially Japan and South Korea. Wilkins' (2004) report based on TIMSS 1995 also showed that the lowest scoring countries on both math and science self-concept were Japan, South Korea and Hong Kong-China. Thus it appears that high self-doubt is a consistent finding in Confucian countries. Lower self-concept means higher anxiety, lower interest and lower enjoyment in science.

The findings from this paper indicate that higher quality of human capital and more R&D inputs are the two magic weapons to national competence in S&T, especially higher quality of human capital. United States lead the world in science and technology generally benefits from the influx of foreign S&T students and highly skilled labor to immigrate to the United States. In order to get higher quality of human capital, two methods offered:

1. Improve K-12 education in general, and science and technology education in particular.

2. Increase the R&D expenditure in order to attract and recruit more R&D researchers and technicians.

REFERENCES

Brody, N. (1997). Intelligence, schooling, and society. *American Psychologist, 52*(10)*,* 1046-1050.

Brown, G., Micklewright, J., Schnepf, S. V., & Waldmann, R. (2007). International surveys of educational achievement: how robust are the findings?. *Journal of the Royal statistical society: series A (statistics in society), 170*(3)*,* 623-646.

Chen, S. S., & Luoh, M. C. (2010). Are mathematics and science test scores good indicators of labor-force quality?. *Social indicators research, 96*(1)*,* 133-143.

Deary, I. J., Strand, S., Smith, P., & Fernandes, C. (2007). Intelligence and educational achievement. *Intelligence, 35*(1)*,* 13-21.

Goolsbee, A. (1998). *Does government R&D policy mainly benefit scientists and engineers?* (No. w6532). National Bureau of Economic Research.

Hall, R. E., & Jones, C. I. (1999). Why do some countries produce so much more output per worker than others?. *Quarterly Journal of Economics, 114*(1)*,* 83-116.

Jensen, A. R. (1980). *Bias in mental testing*. New York: Free Press.

Jencks, C., Smith, M., Acland, H., Bane, M. J., Cohen, D., Gintis, H., et al.(1972).

Inequality: A reassessment of the effects of family and schooling in America. New York: Basic Books.

Lee, J. W., & Barro, R. J. (2001). Schooling quality in a cross–section of countries. *Economica, 68*(272)*,* 465-488.

Lynn, R., & Meisenberg, G. (2010). National IQs calculated and validated for 108 nations. *Intelligence, 38*(4)*,* 353-360.

Mackintosh, N., & Mackintosh, N. J. (2011). *IQ and human intelligence*. Oxford University Press.

Meisenberg, G., & Lynn, R. (2011). Intelligence: A measure of human capital in nations. *Journal of Social, Political and Economic Studies, 36*(4)*,* 421-454.

Sattler, J. M. (2002). *Assessment of children: Behavioral and clinical applications*. Jerome M Sattler Publisher.

Tucker, M. S. (2011). *Surpassing Shanghai: an Agenda for American Education Built on the World's Leading Systems*. Cambridge, MA : Harvard Education Press.

OECD. (2004). *Learning for Tomorrow's World: First Results from PISA 2003.* Paris: OECD.

OECD. (2007). *PISA 2006: Science competencies for tomorrow's world*. Paris: OECD.

Wilkins, J. L. M. (2004). Mathematics and science self-concept: An international investigation. *The Journal of Experimental Education,* 72*,* 331–346.

Wu, M. (2009). A comparison of PISA and TIMSS 2003 achievement results in mathematics. *Prospects, 39*(1)*,* 33-46.

◈◇◈

APPENDIX

Country Name	Article	Population	TIMSS	PISA	IQ	Reseacher	Expenditure
Albania	8	3.2		391			
Argentina	3655	40.3		401	96.0	1092	0.60
Australia	18932	21.3	519	527	98.0		
Austria	4833	8.4		494	99.5	4141	2.71
Azerbaijan	151	8.8		373			0.26
Belgium	7222	10.6		507	99.0	3519	1.96
Brazil	12307	193.7		405	87.0	667	1.17
Bulgaria	735	7.5		439	92.5	1607	0.53
Canada	29017	33.6		529	100.0	4451	1.95
Chile	1868	17	461	447	91.0	286	0.41
Colombia	608	45.7		402	83.5	164	0.16
Croatia	1164	4.4		486	99.0	1593	0.83
Czech Republic	3949	10.4		500	98.0	2743	1.53
Denmark	5307	5.5		499	98.0	6659	3.02
Estonia	518	1.3		528	99.0	3311	1.44
Finland	4952	5.3	552	554	97.0	7644	3.96
France	31757	62.3		498	98.0	3727	2.23
Germany	45017	82.2		520	99.0	3780	2.82
Greece	4882	11.2		470	92.0		
Hungary	2399	10	522	503	96.5	2000	1.15
Iceland	260	0.3		496	101.0	7983	2.82
Indonesia	262	230	406	383	87.0	90	0.08
Ireland	2800	4.5		508	92.5	3217	1.77
Israel	6306	7.2	516	455	95.0		4.27
Italy	26770	59.9	501	489	97.0	1691	1.27
Japan	49632	127.2	558	539	105.0	5147	3.36
Jordan	383	6.3	449	415	84.0		
Kazakhstan	99	15.6	490	400			0.23
Korea, Rep.	22280	48.3	560	538	106.0	5068	3.56
Kyrgyz Republic	15	5.5		330			0.16
Latvia	162	2.2		494		1714	0.46
Lithuania	388	3.3	514	491	92.0	2737	0.84
Luxembourg	137	0.5		484		4811	1.68
Mexico	4128	109.6		416	88.0	369	0.43
Montenegro	11	0.6		401			
Netherlands	14868	16.6		522	100.0	2835	1.84
New Zealand	3188	4.3	512	532	99.0	3724	1.28
Norway	4440	4.8	494	500	100.0	5433	1.80
Panama	73	3.5		376		109	0.21
Peru	159	29.2		369	85.0		
Poland	7359	38.1		508	95.0	1600	0.68
Portugal	4157	10.7		493	94.5	4166	1.66
Qatar	64	1.4	419	379	83.0		
Romania	1367	21.3	465	428	91.0	879	0.48
Russian Federation	14057	140.9	542	478	96.5	3078	1.25

Poland	7359	38.1		508	95.0	1600	0.68
Portugal	4157	10.7		493	94.5	4166	1.66
Qatar	64	1.4	419	379	83.0		
Romania	1367	21.3	465	428	91.0	879	0.48
Russian Federation	14057	140.9	542	478	96.5	3078	1.25
Serbia	1173	9.9		443	88.5	1076	0.89
Singapore	4188	4.7	590	542	108.5	6150	2.20
Slovak Republic	1000	5.4		490	98.0	2450	0.48
Slovenia	1235	2	543	512	96.0	3642	1.86
Spain	21548	44.9		488	97.0	2924	1.38
Sweden	9480	9.2	509	495	99.0	5065	3.62
Switzerland	9472	7.6		517	101.0		
Thailand	2033	67.8	451	425	88.0	332	0.25
Trinidad and Tobago	48	1.3		410			0.06
Tunisia	1022	10.3	439	401	84.0		1.10
Turkey	8307	74.8	483	454	88.5	811	0.85
United Kingdom	45689	61.6	533	514	100.0	4151	1.87
United States	208601	314.7	525	502	98.0	4042	2.82
Uruguay	246	3.4		427	96.0	481	0.44