Revista de Psicodidáctica, 2020, 25 (1), 1–12

Revista de Psicodidáctica

www.elsevier.es/psicod

Original

Exploring the relationship between attitudes toward science and PISA scientific performance☆

Gorka Bidegain[a,*] and José Francisco Lukas Mujika[b]

[a] *Department of Applied Mathematics, University of the Basque Country, Bilbao, Spain*
[b] *Department of Research Methods and Diagnostics in Education, University of the Basque Country, Donostia, Spain*

ARTICLE INFO

Article history:
Received 22 November 2018
Accepted 7 August 2019
Available online 26 September 2019

Keywords:
PISA
Science performance
Attitudes
Quantile regression

ABSTRACT

The Program for International Student Assessment (PISA) 2006 and PISA 2015 are focused on students' competency in science, providing wide data banks for the analysis of the interaction between science performance and attitudes toward science. The few attempts to study this relationship in other assessment studies suggest some positive correlations on the individual level and some unexpected negative correlations and a lack of scalar invariance across countries. The aim of this study is to contribute to the exploration of the generalizability of this relationship across countries and regions within nations. For this, the PISA 2015 data are analyzed using Ordinary Least Square and Quantile regression modelling techniques together with bivariate correlation matrix analysis. The relationship patterns between attitudes—such as self-efficacy, interest in science, participation in science activities, and enjoyment of science—and performance in science are explored at different scales; across 72 PISA participating countries and across 17 regions in Spain. Across countries, the relationship is unexpectedly negative for all attitudes, although high quantiles show a much less pronounced pattern. Across regions, only self-efficacy is significantly and positively correlated with science performance. Overall, positive non–linear relationships are distinguished for high performance values. The results of this study suggest the need of further research using non-parametric quantile regression modeling, and exploring attitudinal indices scaling when investigating potential universal/invariant models. This research should try to justify the comparison across countries/regions using aggregated scores, while incorporating differences in cultural, educational, and social influences on attitudes toward science.

Exploración de la relación entre las actitudes ante las ciencias y el rendimiento en PISA

Palabras clave:
PISA
Rendimiento en ciencia
Actitudes
Regresión cuantílica

RESUMEN

El Programa Internacional para la Evaluación de Estudiantes (PISA) 2006 y PISA 2015 se centran en la competencia científica de los estudiantes, proporcionando amplios bancos de datos para el análisis de la interacción entre el rendimiento de los estudiantes en ciencias y las actitudes no cognitivas hacia la ciencia. Los pocos estudios que exploran esta relación sugieren correlaciones positivas a nivel individual y correlaciones negativas no esperadas junto con una falta de invariabilidad escalar a nivel internacional. El objetivo de este estudio es el de contribuir a la exploración de modelos para la generalización de esta relación entre países y regiones dentro de las naciones. Para ello, se analizan los datos de PISA 2015 utilizando modelos de regresión de Mínimos Cuadrados Ordinarios y de regresión cuantílica, junto con el análisis de la matriz de correlación bivariada. Los patrones de relación entre las actitudes no cognitivas—como la autoeficacia, el interés por la ciencia, la participación en actividades científicas y el disfrute de la ciencia—y el rendimiento en ciencia se exploran a diferentes escalas; 72 países participantes en el PISA y 17 regiones de España. A nivel internacional, la relación es inesperadamente negativa para todas las actitudes, aunque los cuantiles altos (alto rendimiento en ciencia) muestran un patrón mucho

PII of original article:S1136-1034(18)30267-3.

☆ Please cite this article as: Bidegain G, Lukas JF. Explorando la relación entre las actitudes ante las ciencias y el rendimiento en PISA. Rev Psicodidáct. 2020;25:1–12. https://doi.org/10.1016/j.psicod.2019.08.003

* Corresponding author at: Department of Applied Mathematics, University of the Basque Country, Bilbao, Spain.
E-mail address: gorka.bidegain@ehu.eus (G. Bidegain).

menos pronunciado. A nivel regional, sólo la autoeficacia se correlaciona de forma significativa y positiva con el rendimiento científico. En general, las relaciones no lineales positivas se distinguen por sus altos valores de rendimiento. Los resultados de este estudio sugieren la aplicación de modelos no paramétricos de regresión cuantílica y el análisis de las propiedades de los índices de actitud y el efecto de la escala en futuras investigaciones para el desarrollo de modelos universales. Estas investigaciones deberían tratar de justificar la comparación entre países/regiones utilizando puntuaciones medias a la vez que incorporan las diferencias de la influencia cultural, educativa y social en las actitudes hacia la ciencia.

Introduction

The Program for International Student Assessment (PISA) is a standardized comparative study of large-scale international assessments, which is developed in more than 70 participating countries, 34 of them members or associates of the Organization for Economic Co-operation and Development (OECD). According to the OECD (OECD, 2006), science and technology professions are less attractive, and the percentage of students in some OECD countries who are studying science and technology in universities has dropped markedly particularly from 1995 to 2005 (Potvin & Hasni, 2014). The reasons explaining this are varied, but student attitudes toward science may play an important role (OECD, 2006). In consequence, in 2006 and 2015, PISA was focused on students' competency in science, with reading, mathematics, and collaborative problem solving as minor areas of assessment. For this, PISA has incorporated several indices by means of evaluating items in the background questionnaire of its international testing initiative (Krapp & Prenzel, 2011) including students' attitudes toward science (Gallardo et al., 2010).

The most visible results of PISA are the rankings based on aggregated country scores, which have a wide social impact and are being used even to justify reforms in educational legislation (Jornet, 2013). PISA reports are considered by some authors as a stimulating opportunity for seriously thinking about the objective of schools (Gil Pérez & Vilches, 2006; Hernández, 2006). However, there are important concerns about the technical and methodological aspects of validity (Gorur, 2014), comparison between countries (Torney-Purta & Amadeo, 2013), and other issues such as accuracy of measurement of PISA equity (Rutkowski & Rutkowski, 2013). Moreover, there is a lack of explanatory power of assessment studies coupled with the limited use of assessment databases for empirical studies by the scientific community (Pereira, Perales, & Bakieva, 2016).

Regarding science literacy, PISA's extensive data banks are suitable for researching the relationships between science performance and attitudes toward science. This relationship has been previously studied on the individual level, suggesting some positive correlations between, for example, science interest, and self-efficacy and science achievement (Dolin & Evans, 2011). Marsh, Hau, Artelt, Baumert, and Pechar (2006), and Fonseca, Valente, and Conboy (2011) found unexpected negative correlations between achievement and attitudes across countries. Moreover, Täht and Must (2013) found a lack of scalar invariance in educational achievement and learning attitude, suggesting that the relationships between these variables may have a different meaning at the national levels and at the student level within countries. In addition, non-contextualized questions in the student booklet could deliver index values hardly comparable across countries. Overall, this suggests that results based on PISA national average of both test scores and attitudinal scales require a cautious interpretation and a detailed understanding of the indices properties. On the other hand, statistical inferences based on mean scores may include potential etiological influence of aggregated data, distinct from the effects of the same variables on an individual level (Schwartz, 1994).

The question of whether these relationships are universal across different countries, together with the use of PISA national average scores for studying the relationship between learning attitudes and educational achievement, needs further investigation. To explore this relationship, quantile regression can yield a more complete description of functional changes than focusing attention exclusively on the mean (Koenker, 2005) since it is less sensitive to the presence of atypical values.

In the context of a lack of explanatory power of PISA reports and the limited research using assessment databases, the purpose of this study is to contribute to the exploration of the generalizability of the relationship between science achievement and attitudes toward science across and within nations using Ordinary Least Square (OLS) and Quantile regression techniques. The study focuses on self-efficacy in science, interest in broad science topics, participation in science activities, and enjoyment of science.

Method

Participants

The data analyzed in this research were obtained from the PISA 2015 survey test evaluating scientific literacy and background questionnaires. The base PISA target population in each country consisted of 15-year-old students attending educational institutions in grades 7 and higher. National PISA mean scores in science performance and some attitudes toward science from about 540,000 students in 18,618 schools from 72 participating countries and economies were analyzed. In addition, regional data from the 17 autonomous communities (regions) of Spain were similarly analyzed. All data were provided by the OECD (https://www.oecd.org/pisa/data/2015database/) and the Instituto Nacional de Evaluación Educativa (INEE) (http://www.mecd.gob.es/inee/estudios/pisa-2015.html).

Instruments

The survey test for scientific literacy

The test consisted of a computer-based test lasting a total of two hours for each student. Test items were a mixture of multiple-choice questions and questions requiring students to construct their own responses (OECD, 2016a). A set of 185 questions were used in the assessment of science and every student was given only a subset of all questions, and different students receive different sets of questions (OECD, 2016b). Students' scores count toward a national or regional PISA mean score and specific scores in reading, mathematics, and scientific literacy (OECD, 2016a). In this study, the mean scores in scientific literacy of participating countries and regions of Spain were analyzed as response variables (Tables 1 and 2).

Since students work on different test booklets, raw scores were initially scaled so that the OECD average in each domain (mathematics, reading and science) was 500 and the standard deviation was 100 (Stanat et al., 2002). Subsequent PISA cycles are linked

Table 1
National scores (Mean (M) and Standard Error (SE)) of PISA 2015 scientific literacy test and values of indices of attitudes toward science

Country	Science performance		Attitude toward science index							
			SEFISCI		PARTSCI		INTSCI		JOYSCI	
	M	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*
Singapore	556	1.20	.11	.01	.20	.01	.28	.01	.59	.01
Japan	538	2.97	−.46	.02	−.57	.02	−.11	.02	−.33	.02
Estonia	534	2.09	−.04	.02	.29	.02	.02	.01	.16	.01
Chinese Taipei	532	2.69	.19	.02	.20	.01	−.01	.01	−.06	.02
Finland	531	2.39	−.04	.02	−.50	.02	−.09	.02	−.07	.02
Macao	529	1.06	−.03	.02	.17	.02	.06	.01	.20	.01
Canada	528	2.08	.35	.02	−.02	.01	.26	.01	.40	.01
Vietnam	525	3.91	−.28	.03					.65	.02
Hong Kong	523	2.55	−.07	.02	.28	.02	.25	.02	.28	.02
BSJG (China)	518	4.64	−.01	.02	.52	.02	.45	.02	.37	.02
Korea	516	3.13	−.02	.03	−.28	.03	−.07	.02	−.14	.02
New Zealand	513	2.38	−.03	.02	−.20	.02	.09	.02	.20	.02
Slovenia	513	1.32	.07	.02	.07	.02	−.32	.01	−.36	.02
Australia	510	1.54	.07	.01	−.30	.02	.04	.01	.12	.02
Germany	509	2.70	−.01	.02	−.12	.02	.04	.02	−.18	.02
Netherlands	509	2.26	−.08	.02	−.43	.02	−.27	.02	−.52	.02
UK	509	2.56	.27	.02	−.15	.02	.01	.02	.15	.02
Switzerland	506	2.90	−.17	.02	−.12	.02	.15	.02	−.02	.02
Ireland	503	2.39	.06	.02	−.37	.02	.06	.02	.20	.02
Belgium	502	2.29	−.10	.02	−.13	.02	.07	.01	−.03	.02
Denmark	502	2.38	.08	.02	−.13	.02	.18	.02	.12	.02
Poland	501	2.51	.16	.02	.40	.02	−.24	.02	.02	.02
Portugal	501	2.43	.27	.02	.20	.02	.27	.02	.32	.02
Norway	498	2.26	.19	.02	−.04	.02	.05	.02	.12	.02
United States	496	3.18	.26	.02	−.02	.02	.05	.02	.23	.02
Austria	495	2.44	−.17	.02	−.14	.02	.06	.02	−.32	.02
France	495	2.06	−.13	.02	−.11	.02	−.06	.02	−.03	.02
Czech R.	493	2.27	.10	.02	−.08	.02	−.67	.01	−.34	.02
Spain	493	2.07	−.14	.02	−.20	.02	.10	.01	.03	.02
Sweden	493	3.60	.05	.02	−.25	.02	−.02	.02	.08	.03
Latvia	490	1.56	−.01	.02	.22	.02	.14	.01	.09	.02
Russia	487	2.91	.02	.03	.66	.02	.03	.02	.00	.02
Luxembourg	483	1.12	−.03	.02	.07	.02	.21	.01	.10	.02
Italy	481	2.52	.13	.02	.27	.02	.21	.02	.00	.02
Hungary	477	2.42	−.05	.02	.27	.03	−.23	.02	−.23	.02
Argentina	475	6.28	−.04	.05					−.20	.04
Croatia	475	2.45	.10	.02	.03	.02	−.16	.02	−.11	.02
Lithuania	475	2.65	.26	.02	.37	.02	.11	.01	.36	.02
Iceland	473	1.68	.24	.03	−.17	.02	.23	.02	.15	.02
Israel	467	3.44	.04	.02	.09	.04	−.24	.02	.09	.02
Malta	465	1.64	−.09	.02					.18	.02
Slovak Republic	461	2.59	−.06	.02	.14	.02	−.32	.02	−.24	.02
Kazakhstan	456	3.67	.46	.03					.85	.02
Greece	455	3.92	−.04	.02	.19	.02	.14	.02	.13	.02
Chile	447	2.38	−.10	.02	.17	.02	.04	.02	.08	.02
Bulgaria	446	4.35	.39	.02	.82	.02	.28	.02	.28	.02
Malaysia	443	3.00	−.13	.02	.88	.02	.49	.02	.52	.02
United Arab Emir.	437	2.42	.41	.02	.88	.02	.19	.01	.47	.02
Romania	435	3.23	−.20	.02					−.03	.02
Uruguay	435	2.20	.05	.02	.14	.02	−.05	.02	−.10	.02
Cyprus	433	1.38	−.05	.02	.46	.02	.02	.02	.15	.02
Argentina	432	2.87	−.10	.02					−.09	.02
Moldova	428	1.97	.09	.02					.33	.01
Albania	427	3.28	.02	.02					.72	.02
Turkey	425	3.93	.35	.02	.68	.02	−.06	.02	.15	.02
Trinidad Tobago	425	1.41	.11	.02					.19	.02
Thailand	421	2.83	.17	.02	.92	.02	.60	.01	.42	.01
Costa Rica	420	2.07	−.12	.02	.31	.02	.22	.02	.35	.02
Qatar	418	1.00	.36	.02	.80	.01	.25	.01	.36	.01
Mexico	416	2.13	.27	.02	.53	.02	.43	.01	.42	.02
Colombia	416	2.36	−.05	.02	.64	.02	.35	.01	.32	.01
Georgia	411	2.42	.27	.02					.34	.02
Montenegro	411	1.03	.31	.02	.86	.02	−.08	.02	.09	.02
Jordan	409	2.67	.56	.03					.53	.02
Indonesia	403	2.57	−.51	.02					.65	.01
Brazil	401	2.30	.17	.02	.50	.02	.24	.01	.23	.01
Peru	397	2.36	.34	.02	.70	.02	.46	.01	.40	.01
Lebanon	386	3.40	.17	.03					.38	.02
Tunisia	386	2.10	−.07	.02	1.20	.02	.26	.01	.52	.02
FYROM	384	1.25	−.06	.02					.48	.02
Kosovo	378	1.70	−.29	.02					.92	.02
Algeria	376	2.64	−.16	.02					.46	.02
Dominican Rep.	332	2.58	.54	.04	.92	.03	.69	.02	.54	.02

Table 2
Regional scores (Mean (M) and SE) of PISA 2015 scientific literacy test in Spain and values for indices of attitudes toward science

Region	Science	Performance	Attitude toward science index							
			SEFISCI		PARTSCI		INTSCI		JOYSCI	
	M	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*
Castilla León (CL)	519	3.52	−.01	.03	−.20	.03	.19	.02	.11	.04
Madrid (MAD)	516	3.45	−.01	.04	−.20	.04	.13	.02	.13	.03
Galicia (GAL)	512	3.14	−.08	.04	−.30	.04	.05	.03	.16	.03
Navarra (NAV)	512	4.09	−.27	.03	−.40	.03	.03	.02	−.06	.03
Aragón (ARA)	508	4.65	−.13	.03	−.30	.03	.11	.03	−.01	.04
Cataluña (CAT)	504	4.73	−.10	.05	−.10	.03	.09	.03	.06	.04
Asturias (AST)	501	3.88	−.03	.04	−.20	.04	.05	.03	.02	.03
La Rioja (RIO)	498	5.50	−.05	.04	−.20	.04	.07	.03	−.03	.03
Castilla La Mancha (CM)	498	4.02	−.06	.03	−.20	.04	.16	.02	.03	.03
Cantabria (CANT)	496	5.60	−.07	.05	−.30	.03	.08	.03	.01	.04
Comunidad Valenciana (CV)	494	3.25	−.09	.05	−.20	.04	.08	.03	−.03	.04
Islas Baleares (IB)	485	4.52	−.21	.04	−.10	.03	.07	.03	.01	.03
Murcia (MUR)	484	3.79	−.17	.03	−.20	.03	.13	.02	.06	.03
Esukadi (EUS)	483	3.02	−.34	.03	−.30	.03	.04	.02	−.05	.03
Islas Canarias (ICAN)	475	3.56	−.18	.03	−.10	.04	.16	.03	.11	.04
Extremadura (EXT)	474	3.79	−.19	.04	−.20	.03	.08	.02	−.01	.04
Andalucía (AND)	473	4.15	−.23	.06	−.30	.04	.13	.03	−.06	.04

to the previous cycles through item response theory scale linking methods (Mazzeo & von Davier, 2013). The scaling procedures are described in detail in OECD (2014), including different scales' reliabilities (Cronbach's Alpha) for attitudinal indices.

The background questionnaire and attitudinal indices

To gather contextual information, students also answered a background questionnaire based on the context questionnaire framework described in OECD (2016b). In PISA 2015, as in 2006, attention was focused on student attitudes to science both inside and outside the classroom. PISA sought information on attitudinal measures toward science (OECD, 2009). In this study, the PISA attitudinal indices analyzed were: (1) science self-efficacy (SEFISCI); (2) participation in science activities (PARTSCI); (3) general interest in science learning (INTSCI); and (4) enjoyment of science (JOYSCI). The reports on these indices were constructed based on OECD mean of logit scores with equally weighted country samples, so that the average score across OECD countries is zero, and two-thirds score between 1 and −1. A positive value on the index indicates that students report a higher score than the OECD average. A negative value indicates a lower attitude value than the OECD average.

Self-efficacy in science (SEFISCI)

Academic self-efficacy refers to individuals' convictions that they can successfully perform given academic tasks at designated levels (Schunk, 1991). That is, the constructed index of self-efficacy in science (SEFISCI) measures how much students believe in their own abilities to handle tasks and overcome difficulties effectively. Students responded to eight items on a four-point Likert scale, stating whether they felt "very confident," "confident," "not very confident," or "not at all confident" (OECD, 2016a).

Interest in broad science topics (INTSCI)

In order to measure students' general interest in science subjects, PISA asked students a set of questions on: (1) their level of interest in different scientific subjects; (2) their interest in the ways in which scientists design experiments; and (3) their understanding of what is required for scientific explanations (OECD, 2016a). Students declared their interest on a four-point Likert scale with the categories "not interested," "hardly interested," "interested," and "highly interested" (OECD, 2014).

Participation in science activities (PARTSCI)

Another set of questions from PISA addressed students' participation in science activities both at school and out of school, such as science competitions, joining a science club (OECD, 2016a) and watching TV programs, reading magazines, visiting websites, etc. Students were asked how often they engaged in science related activities on a four-point scale with the answering categories "very often," "regularly," "sometimes," and "never or hardly ever" (OECD, 2014).

Enjoyment of science (JOYSCI)

In addition to INTSCI, another aspect that relates to students' motivation to learn science is the enjoyment of science, that is, how interesting and fun students find learning science (OECD, 2016c). Enjoyment of science is an index obtained from a set of questions asking students to respond on a four-point Likert scale with the categories "strongly agree," "agree," "disagree," and "strongly disagree" (OECD, 2014). For example, PISA asked students to indicate their level of agreement with the following statements: (1) I generally have fun when I am learning science topics; (2) I like reading about science; (3) I am happy doing science problems; (4) I enjoy acquiring new knowledge in science; and (5) I am interested in learning about science (OECD, 2012).

Data analysis

First, bivariate correlation analysis between all variables was conducted. The selection of an adequate correlation test (i.e. Pearson's parametric test or Spearman's non-parametric test) was based on the distribution of the variables. The distribution of each variable and significance levels (p-value) and correlation coefficients were computed for both tests by means of a correlation matrix. Non-linear fitted curves were also computed for an improved exploratory analysis.

For a more detailed analysis, the Ordinary Least Squares regression model (OLSrm) was used to investigate relationship patterns between attitude-toward-science variables (predictor variables), and science achievement (response variables). For further analysis of the models, residuals of regressions were checked, and non-linear relationships between variables not explained by the regression models nor influential cases (i.e. influential outliers) were found. Residuals were analyzed after fitting the models to data in case they could reveal unexplained patterns in the data by the fitted model and, consequently, to check if linear regression assumptions were met while improving the models in an exploratory way.

Similarly, Quantile regression models (Qrm) were applied for 5, 50, and 95 quantiles. The Qrm for the attitude indices *vs* science performance relationship were applied to the across-country case. The sample size (17 regions) for the across-regions case was considered too small for an adequate quantile regression analysis. The equation $PISA\ Science\ Score = a + b \times Attitude\ Index$ represents the regression model for the analyzed relationship, where coefficient *a* represents the intercept and coefficient *b* represents the slope. These coefficients were used to compare the OLSrm with the Qrm. Qrm estimates were compared with OLSrm at the intercept and slope of the models. Firstly, for comparing intercepts between models, the confidence interval test (CI = 95%) was performed (du Prel, Hommel, Rohrig, & Blettner, 2009), which assessed the pairwise differences between sites. The differences were considered significant ($p < .05$) when the confidence interval did not contain zero. Secondly, a Quantile Regression Analysis of Deviance was performed to test the equality of the slopes of the models. These analyses were performed across PISA participating countries (OLSrm, Qrm) and across regions within Spain (OLSrm). The quantile analysis was not performed on the regional data due to the limited number of regions (17 regions). This study did not explore the effect of sex in the analysis since males and females do not have significantly different attitudes toward school science (OECD, 2009).

The aforementioned analyses were conducted with the software R (R Core Team, 2017). In particular, the following packages were used: rgl, quantreg (Koenker, 2017), car, corrplot, ggplot2, multiplot, visreg, devtools, easyGgplot2, SparseM, Rmisc, grid, gridExtra, and Performance Analytics.

Interpretation of statistical data

The correlations and regression models obtained in this study from mean scores need cautious interpretation. Ecological inference fallacy (Ess & Sudweeks, 2001) occurs when inferences about the nature of individuals are deduced from inferences about the group to which those individuals belong. Averaging may reduce scatter and increases correlation. The use of mean scores as data points may overestimate the correlation between the student science performance and attitudes toward science. Ecological variables are necessary to examine structural, contextual, and sociological effects on human behavior (Schwartz, 1994). The relationship between attitudes toward science and performance in science may be an example of this. Attitudes are not reducible to individual characteristics. Sociocultural environment has a potential influence on the attitudes toward science. Here, correlations and regression models based on mean scores tries to take note

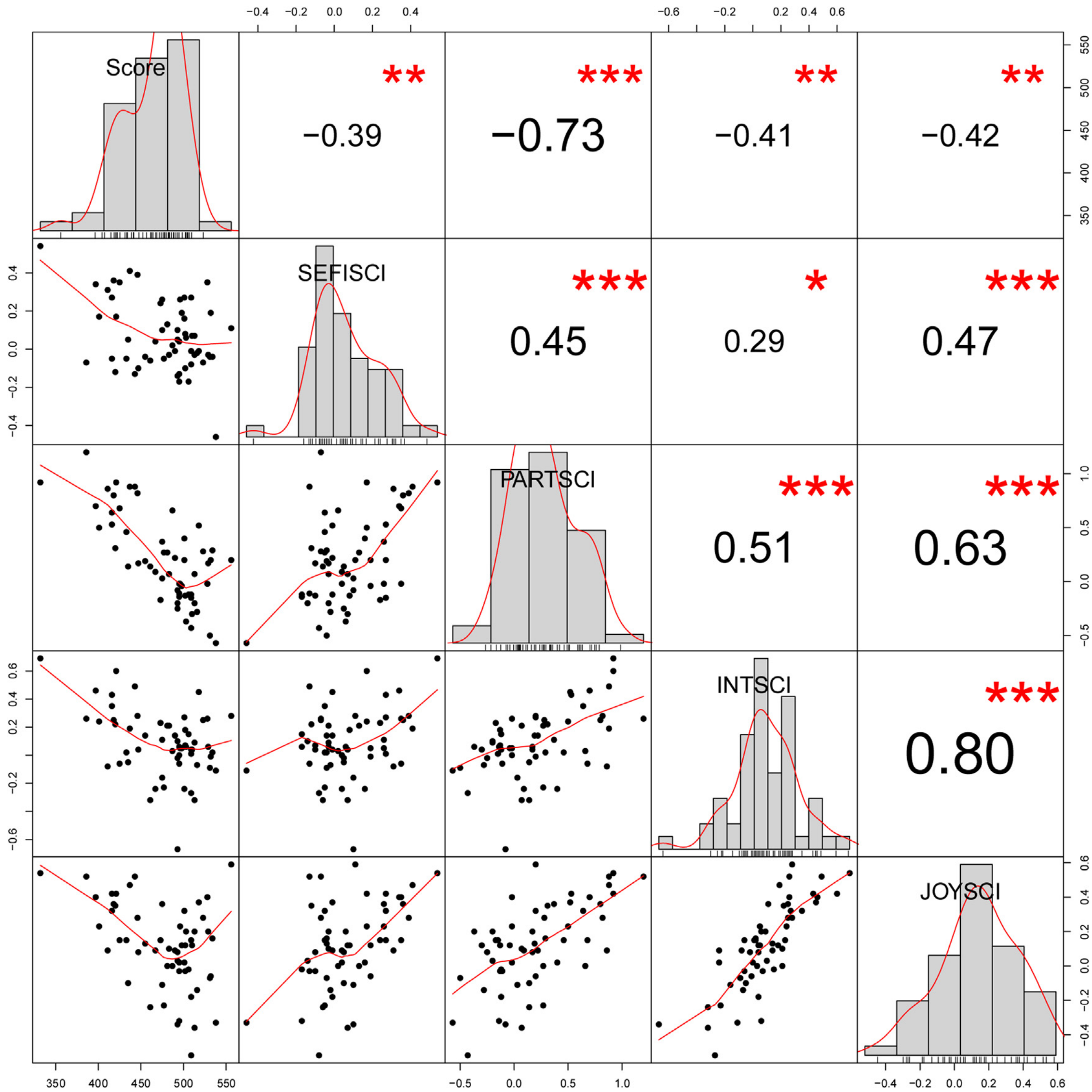

Figure 1. Correlation matrix (countries) for the relationship between the 'attitude toward science' variables and the performance in science literacy (PISA 2015 score) across participating countries. The distribution of each variable is shown on the diagonal. On the bottom of the diagonal, the bivariate scatter plots with a fitted non-linear curve are displayed. On the top of the diagonal the value of the correlation plus the significance level as stars; p-values .001, .01, .05, .1 for symbols "***", "**", "*", "".

of the potential etiological influence of aggregate-level variables, distinct from the effects of the same measures on an individual level (Schwartz, 1994).

Results

Cross-country analysis

Correlations among the studied attitude indices were significant and overall showed moderate and strong positive correlations (Figure 1). JOYSCI and INTSCI were the two indices with the highest degree of correlation ($r = .80$). Correlations between the attitude variables and science achievement scores showed significant negative correlations, particularly strong for PARTSCI ($r = -.73$). Additionally, the non-linear fitted curves (diagonal scatterplots in Figure 1) showed a pattern of non-negative or positive relationship between the explored indices and performance at the upper extremes of the curves (i.e., for high science performance scores).

The OLS regression models (Figure 2) showed significant negative linear relationships ($p < .05$) between attitude indices and science performance, with intercepts between 467 (SEFISCI) and 493 (PARTSCI) and slopes between −41 (SEFISCI) and −88 (JOYSCI). Interestingly, countries with high science performance scores such as Japan (538) and Finland (531) obtained low and very low values for all SEFISCI indices (Table 1). On the other side, countries like Turkey, Mexico, and the Dominican Republic with low and very low science performance scores obtained generally high values for all attitudinal indices (Table 1).

Similarly to the OLSrm, Q50 (the median) showed negative relationships for all indices (Figure 3). However, as the quantile was increased, the negative slope almost disappeared for all indices. When the Q5 quantile was analyzed, the relationships for both PARTSCI and INTSCI with performance became even more negative, while for SEFISCI, the relationship became slightly positive, and for JOYSCI, it did not change significantly.

Qrm coefficients were compared to OLSrm coefficients (intercept and slope) along quantile values from Q10 to Q90, as shown in Figure 4. For all indices, the quantile regression point intercepts (coefficient a) lay outside (above) the OLSrm confidence

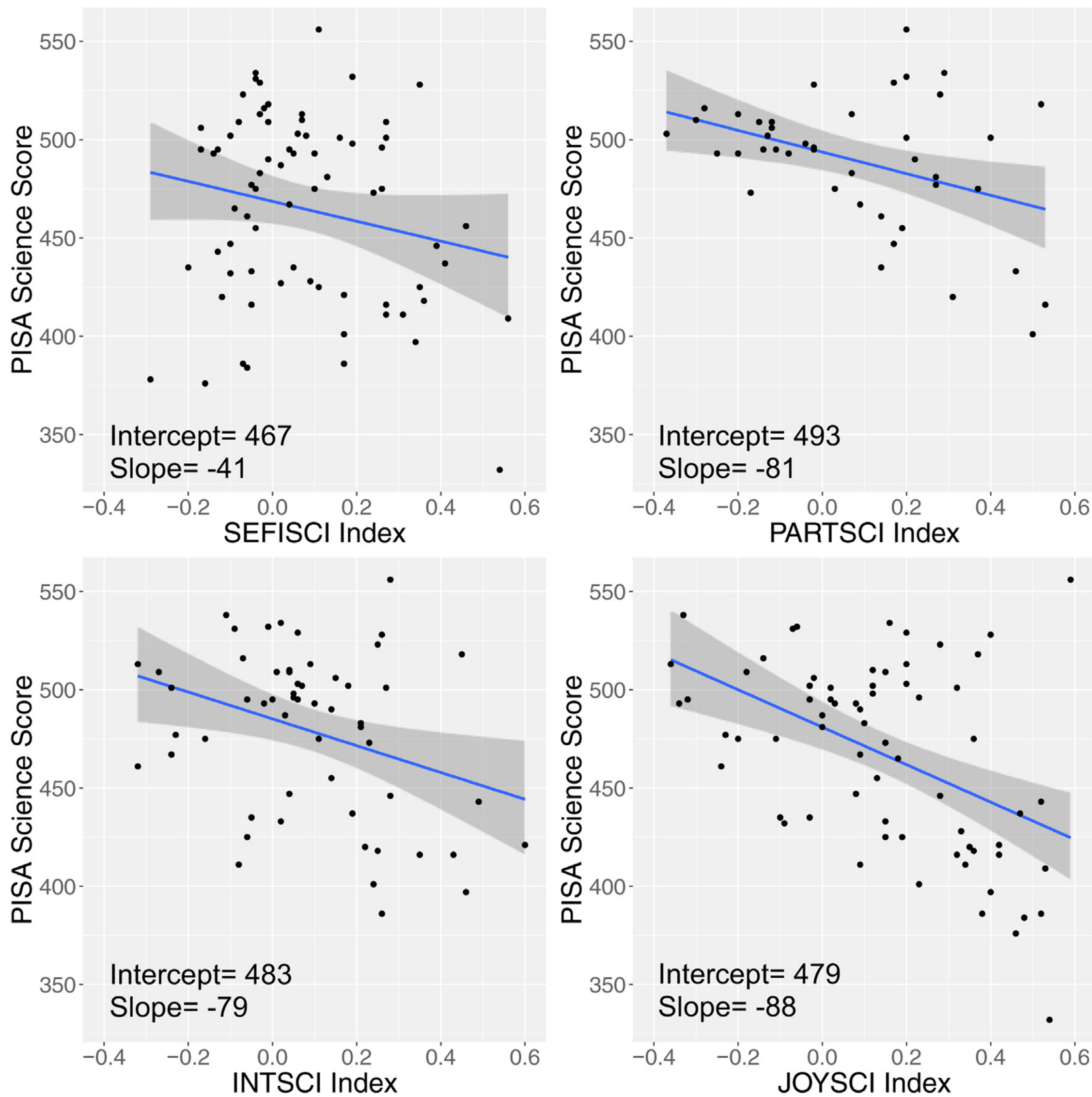

Figure 2. OLSrm for the relationship between variables related with attitudes toward science and performance in scientific literacy in PISA participating countries. The shaded grey area form 95% point-wise confidence band around the OLSrm estimates (blue line).

interval beyond the quantiles Q60 or Q70 (see top panels of each subfigure A, B, C, and D of Figure 4). For high quantiles, the intercept (i.e., performance score at the point where the value for the attitudinal index is zero; that is, the average index score across OECD countries) was significantly higher ($p < .05$) than for low quantiles.

Regarding the slope or the estimate of the models (coefficient *b*), the quantile regression point estimates and the OLSrm curves were not statistically different along the different quantile values for SEFISCI ($F = .53$, $p = .83$) and INTSCI ($F = .88$, $p = .54$) when tested through the Quantile Regression Analysis of Deviance. For these two indices, the quantile regression curve (including the confidence interval) of coefficient *b* clearly overlapped the OLSrm confidence interval (see bottom panels of subfigures A and C of Figure 4).

For PARTSCI and JOYSCI, the quantile regression curve of coefficient *b* lays above the OLSrm line beyond quantile Q70 (see bottom panels of subfigures B and D of Figure 4). Although the confidence intervals of both models overlapped, the Quantile Regression Analysis of Deviance confirm significantly higher estimates for Qrm than for OLSrm for high quantiles for both PARTSCI ($F = 2.09$, $p = .035$) and JOYSCI ($F = 3.43$, $p = .0007$). This means that for high quantile values, the relationships for both PARTSCI and JOYSCI with performance are not as negative as for the mean based linear model resulting in coefficients (slopes) near zero.

Regional analysis: the case of Spain

For within-country regional analysis (Table 2), after checking normality assumption, the Spearman's correlation test was applied. The correlation analysis among the studied attitude indices showed a moderate and significant correlation ($r = .57$) only for SEFISCI *vs* JOYSCI (Figure 5). Attitude variables and science achievement scores showed a significant positive correlation for SEFISCI ($r = .61$). The non-negative or positive relationship pattern between the explored indices and performance at the upper extremes of the curves can be also observed for the regional analysis in Spain (see fitted non-linear curves in the diagonal plots of Figure 5).

SEFISCI showed a significant positive linear relationship ($p < .01$) with performance with an intercept value of 507 and a slope of 91 (Figure 6). That is, regions with low SEFISCI values such as Euskadi (the lowest SEFISCI value) and Andalucia, for example, obtained significantly lower science performance scores than Castilla León or Madrid, which has the highest SEFISCI values (Table 2). The rest

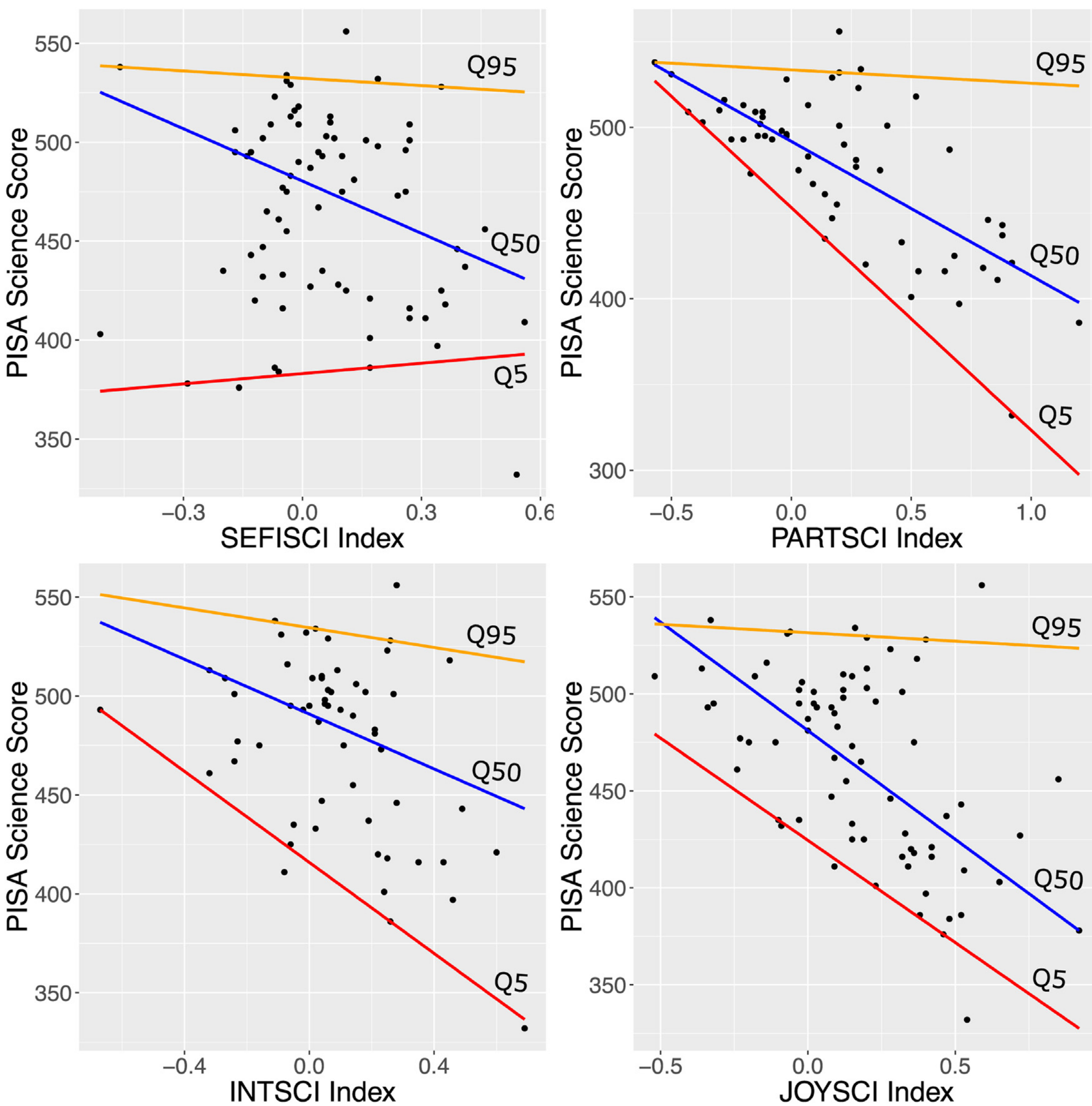

Figure 3. Quantile regression curves for attitude variables (SEFISCI, PARTSCI, INTSCI, and JOYSCI) and PISA score in science across PISA 2015 participating countries using linear quantile regression techniques (Koenker, 2005, 2017) are represented as follows: quantile 5 (Q5, red line), quantile 50 or median (Q50, blue line) and quantile 95 (Q95, orange line).

of the regression models were not significant as can be noted in the case of JOYSCI *vs* Science Performance with a $p = .1$.

The quantile regression models were not applied to the cross-country case because of the size of the small size of the sample for an adequate analysis.

Discussion

A universal understanding of the relationship between educational attitudes and science performance is essential to improve teaching and learning systems and to better understand the decline in the share of science students in universities across the world in previous decades (Fonseca et al., 2011; Newell, Tharp, Vogt, Moreno, & Zientek, 2015). General and transferable models relating attitudes toward science and science performance were not obtained in this study. However, some interesting insights were gained in this regard. The results, discussed below, suggest that quantile analysis, together with non-linear or non-parametric regression models may be an interesting approach for further research when investigating potential universal (invariant) models using nationally and regionally aggregated scores. Moreover, this type of study demonstrates the extent to which comparisons between countries are always stimulating; multiple perspectives are needed to assess the situation of each country and how such vision, both national and regional, may be enriched when placed within a broader framework for a more general interpretation of the science performance scores (INEE, 2017).

The cross-country analysis showed significant correlations among the studied attitude indices, with particularly strong correlations for JOYSCI and INTSCI. Although the constructs are independent, they may be related to each other (Baram-Tsabari & Yarden, 2009). Krapp and Prenzel (2011) suggest that these two attitudes may overlap because enjoyment "can occur for many reasons, and interest is only one of them." The PISA 2015 (OECD, 2016a) definition for interest includes "curiosity" and "willingness." This link between the indices and their diverse definitions may suggest the need for further research into the definitions of the attitudinal indices and assessment methods.

Positive attitudes toward science may improve students' academic performance in science as resulted from individual level

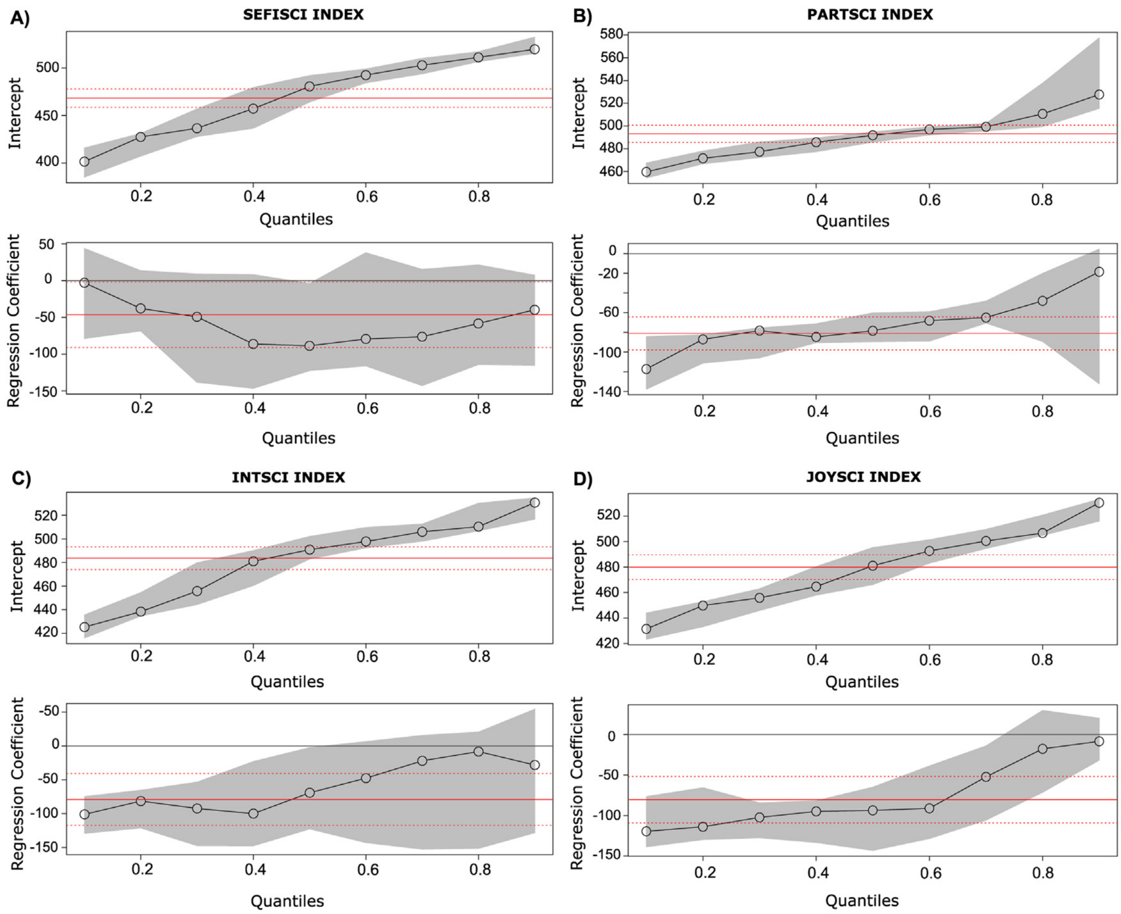

Figure 4. Comparison between the quantile regression model point estimates (black line with circles) with the OLSrm results (red line) at the intercept and regression coefficient (slope) for regression models obtained for the relationship between the attitude toward science variables and performance in scientific literacy in PISA 2015 participating countries. The 95% confindence interval for OLSrm and Qrm points is represented by red dash lines and grey shaded area, respectively.

studies (Dolin & Evans, 2011). However, on a broader scale (i.e., across nations and regions), the relationship between all attitude variables and science achievement using aggregated PISA scores was found to be negative (Figures 1 and 3), and particularly strong for PARTSCI. This result is in line with Marsh et al. (2006) and Fonseca et al. (2011) who found unexpected negative correlations between educational achievement and attitudes at the national level in some countries.

Countries with high PISA scores in science, where schools reported to provide few science activities and showed relatively low PARTSCI index values, are, among others, Japan (−.57), Finland (−.60), Netherlands (−.43), and Korea (−.28) (Table 1). On the contrary, some countries with high PARTSCI index values such as Tunisia (1.20), Thailand (.92), Montenegro (.86), and Peru (.70) showed high science performance scores (Table 2) (OECD, 2016a). The absence of a clear pattern for the relationship between PARTSCI and science performance may have two possible explanations, considering the fact that the studied relationships are not causal. On the one hand, students may develop alternative ideas about science from activities such as watching TV programs, reading magazines, visiting websites, affecting their performance in science (Chi, 2000). Alternative conceptions tend to be very resistant to instruction (Chi, 2005). On the other hand, the difficulties faced by PISA in gathering reliable data and scaling data about student attitudes by posing non-contextualized questions in the student questionnaire could result in index values that cannot be compared across-countries yet, since they may not incorporate students' broader historical and cultural context. Overall, this suggests that results based on PISA national average of both test scores and attitudinal scales require a cautious interpretation and a detailed understanding of the indices.

The non-linear fitted curves (Figure 1, diagonal) showed a pattern of non-negative or positive relationship between the explored indices and performance at the upper extreme of these curves (i.e., for high science performance score) as reported by Fonseca et al. (2011). Overall, the Qrm applied to the international data (Figure 3) showed that as the quantile was increased the negative slope almost disappeared for all indices. In addition, the comparison of Qrm coefficients to OLSrm coefficients along quantile values (Figure 4) showed differences in the relationship between attitudes and performance. Particularly for PARTSCI and JOYSCI, the relationship between the indices and performance found with the Qrm was not as negative as that found with the mean based linear model (OLSrm). For high quantiles, Qrm coefficients (slopes) were near zero; that is, non-significant relationships between indices and performance (Figure 4).

The results demonstrated the suitability of the quantile regression models to record specific behaviors in the interaction of attitudes toward science and science performance as an alternative to mean-based models. However, a full interpretation of these results may require further research. The preliminary results of the fitted non-linear curves (Figures 1 and 5, scatter plots) suggest exploring the use of non-parametric simple regression models and non-parametric quantile regression models (Lipsitz, Belloni, Chernozhukov, & Fernández-Val, 2016) to provide a more accurate estimate of the regression function

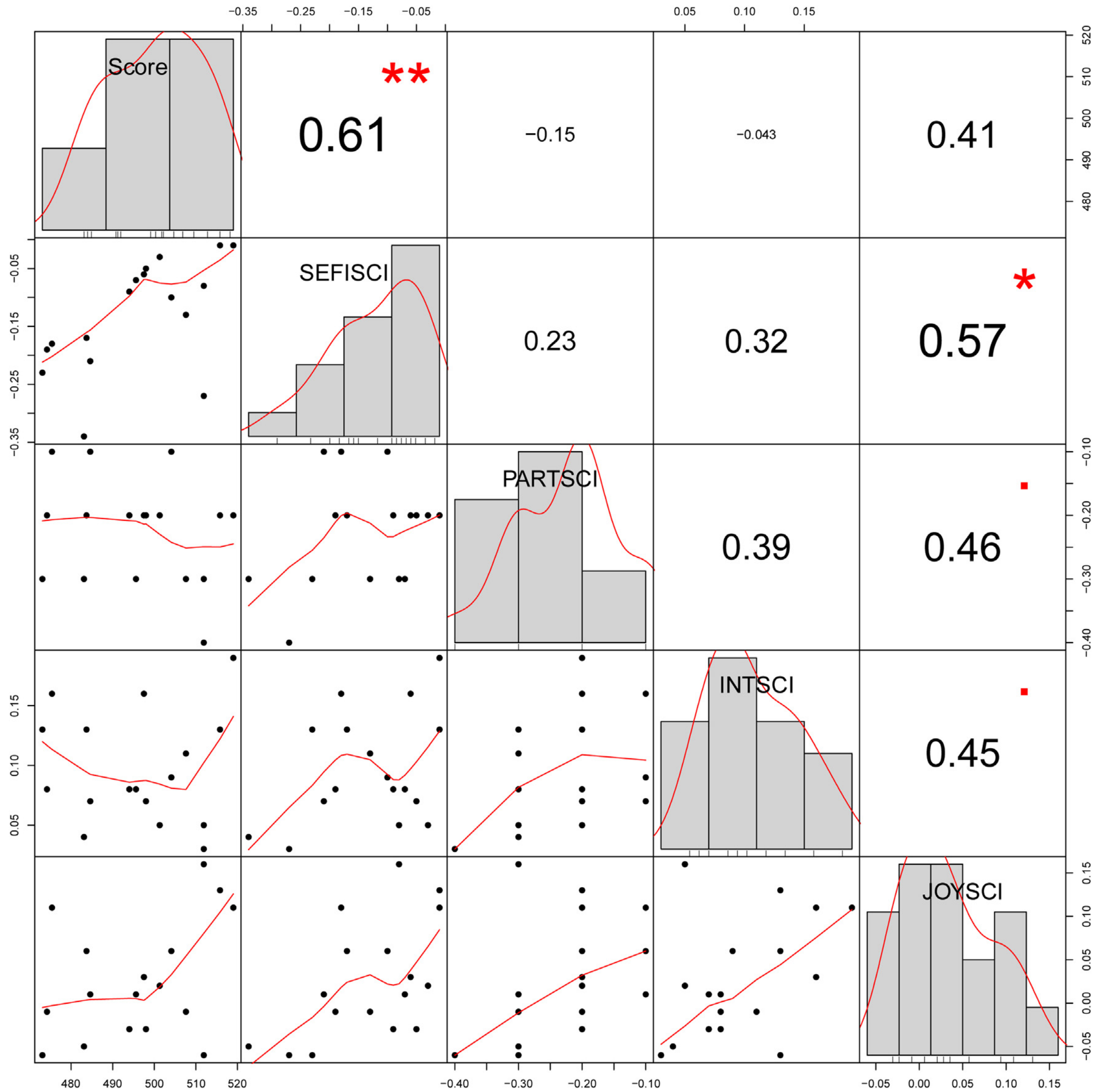

Figure 5. Correlaton matrix (regions, Spain) for the relationship beween the ‘attitude toward science’ variables and the performance in science literacy (PISA 2015 score) across regions in Spain. On the top of the diagonal the value of the correlation plus the significance level as stars; p-values .001, .01, .05, .1 for symbols “***”, “**”, “*”, “”.

than that obtained by the parametric analysis applied in this study.

The cross-region analysis showed non-significant relationships between attitudinal indices, except for SEFISCI vs JOYSCI, which indicated a moderate positive correlation (Figure 5). The term self-efficacy in science is used to describe students' positive belief in solving a difficult scientific problem, and it is a powerful incentive to persevere in spite of difficulties (Bandura, 1977) and may be related to enjoyment learning science. Regarding science performance, regional high SEFISCI index values were significantly correlated with PISA score in science (Figures 5 and 6), but the remaining indices showed non-significant association with science performance. Self-efficacy in science has been related to performance, their career orientation, and their choice of courses (Nugent et al., 2015). Considering the OECD data, students' self-efficacy beliefs have shown to be strongly correlated with student achievement both within and between some countries (OECD, 2016c). It is also interesting to note the feedback effect suggested by Bandura (1997), where better academic results in science may lead to an increase in self-efficacy through positive feedback received from teachers, peers, and parents.

The cross-region analysis within countries might be affected by its limited data (17 regions in the present case) and this effect should be further assessed. Thus, caution is needed when interpreting the cross-regional results of this study. Regions with low SEFISCI values such as Euskadi and Andalusia obtained significantly lower science performance scores than Castile León or Madrid, which had the highest SEFISCI values (Figure 6, Table 2). The case of Euskadi, for example, is remarkable. Here, the regional gross domestic product (GDP) is about €70,000/month, far above the national and EU28 average (Eurostat, 2018). This autonomous community scored relatively low in all indices and in science performance (14th out of 17; 9th considering significant differences) (Table 2). This result contradicts the fact that socio-economically advantaged countries and regions may be considerably more likely to have students with good science attitudes and PISA scores (OECD, 2016a).

It is necessary to understand that this and other relationships linked to family background, socio-economic status, and immigration background (Willms, 2006) may have become myths (Schleicher, 2016). Over the last years (1) less than a quarter of the performance variation among OECD countries is explained by gross domestic product (GDP) per capita, and (2) many socio-economically disadvantaged countries and regions have changed

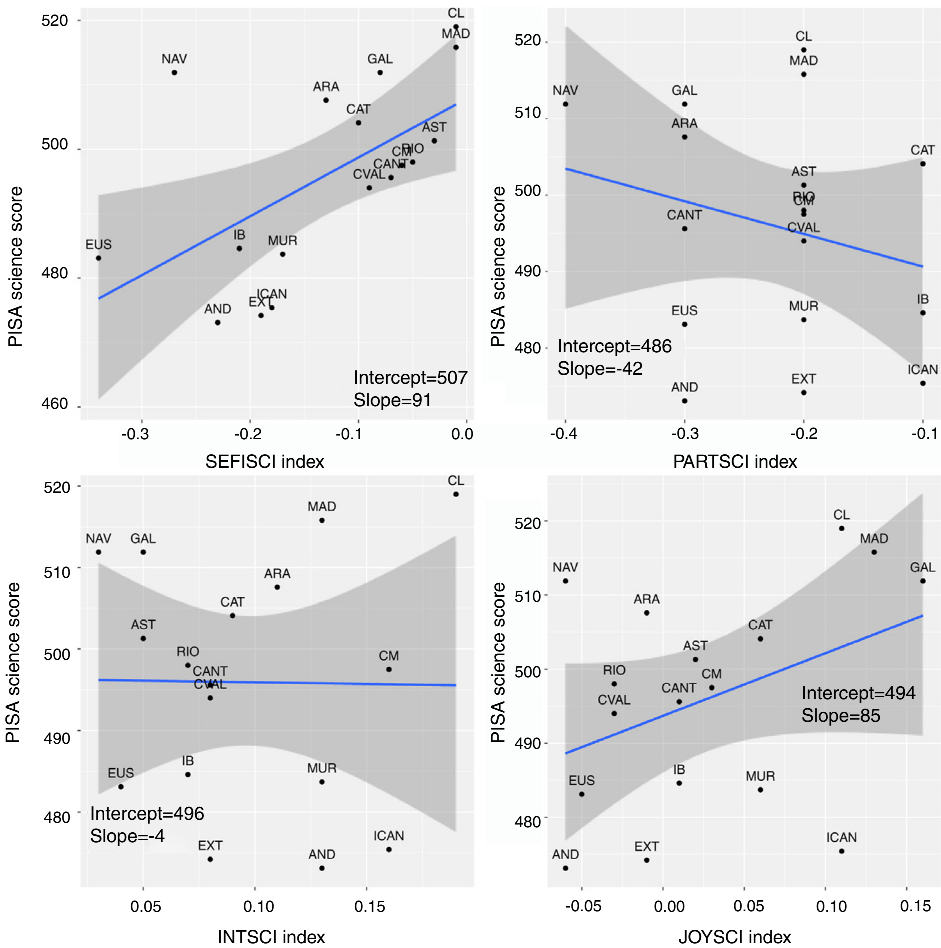

Figure 6. OLSrm for the relationship between variables related with the attitude toward science and performance in scientific literacy (PISA) in the Autonomous Communities of Spain. The shaded grey area form 95% point-wise confidence band around the OLSrm estimates (blue line).

their education policies and practices, leading to significant gains in learning outcomes. As shown in science literacy (Tables 1 and 2), there are some countries and regions that are not wealthy but provide excellent education, and countries and regions that are significantly developed but are not competitive in PISA rankings. An example that education and public policy can make a great difference for students that are disadvantaged is the fact that ten percent of the most disadvantaged students in Shanghai obtained similar science scores to those obtained by 10 percent of the most privileged American fifteen-year-olds (OECD, 2016a). Highly developed countries and regions with low performance and low scores in science attitudes may have to take note of some factors, such as the quality of teachers or the benefit of a positive feedback has on students' attitudes. Highest-performing education systems prioritize the quality of teachers over the size of classes (Schleicher, 2016).

Highly developed countries and regions with low scores in science performance, may be critical of their educational system. However, this self-criticism need not lead to a blind path towards the sole improvement of PISA test scores. PISA may be more interested in following the laws and scientific theories through closed questions restricted to problems of a purely academic nature, with little or no connections between disciplines, rather than encouraging the divergent thinking implicit in the creative spirit, which is fundamental to the advancement of science (Gallardo et al., 2010).

In conclusion, the results of this study suggest that further research using non-parametric quantile regression curves and scaling may be critical when investigating potential universal/invariant models across countries and regions. This research may try to justify the comparisons across countries with nationally aggregated scores. Correlations and regression models based on mean scores may take note of the potential etiological influence of aggregate data. This must be done by bearing in mind or incorporating when possible the differences in cultural, educational, and social influences on attitudes toward science.

Acknowledgements

This investigation was conducted under the framework of the Master in Teacher for Training for Compulsory Secondary Education, Upper Secondary Education, Vocational Training, and Foreign Language Teaching of the University of Basque Country. We appreciate this support.

References

Bandura, A. (1977). *Social learning theory*. New York: General Learning Press.
Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.
Baram-Tsabari, A., & Yarden, A. (2009). Identifying meta-clusters of students' interest in science and their change with age. *Journal of Research in Science Teaching*, *46*(9), 999–1022. http://dx.doi.org/10.1002/tea.20294
Chi, M. T. H. (2000). Self-explaining: The dual processes of generating inference and repairing mental models. In R. Glaser (Ed.), *Advances in instructional psychology (vol 5): Educational design and cognitive science* (pp. 161–238). Mahwah: Erlbaum.
Chi, M. T. H. (2005). Commonsense conceptions of emergent processes: Why some misconceptions are robust. *Journal of the Learning Sciences*, *14*, 161–199. http://dx.doi.org/10.1207/s15327809jls1402_1
Dolin, J., & Evans, R. H. (2011). Initial teacher education and continuing professional development for science teachers: A review of recent research results.

In A. Delhaxhe (Ed.), *Science education in Europe: National policies, practices and research* (pp. 103–110). Brussels: Education, Audiovisual and Culture Executive Agency. http://dx.doi.org/10.2797/7170

du Prel, J. B., Hommel, G., Rohrig, B., & Blettner, M. (2009). Confidence interval or p-value?: Part 4 of a series on evaluation of scientific publications. *Deutsches Ärzteblatt International*, *106*(19), 335–339. http://dx.doi.org/10.3238/arztebl.2009.0335

Ess, C., & Sudweeks, F. (2001). *Culture, technology, communication: Towards an intercultural global village*. Albany, New York: State University of New York Press.

Eurostat. (2018). *European commission eurostat database*.

Fonseca, J., Valente, M. O., & Conboy, J. (2011). Student characteristics and PISA science performance: Portugal in cross-national comparison. *Procedia - Social and Behavioral Sciences*, *12*, 322–329. http://dx.doi.org/10.1016/j.sbspro.2011.02.041

Gallardo, M., Fernández-Navas, M., Sepúlveda-Ruiz, M. P., Serván, M. J., Yus, R., & Barquín, J. (2010). PISA y la competencia científica: Un análisis de las pruebas de PISA en el Área de Ciencias. *Relieve*, *16*(2), 1–17.

Gil Pérez, D., & Vilches, A. (2006). ¿Cómo puede contribuir el proyecto PISA a la mejora de la enseñanza de las ciencias (y de otras áreas de conocimiento)? *Revista de Educación, Volumen Extra PISA*, 295–331.

Gorur, R. (2014). Towards a sociology of measurement in education policy. *European Educational Research Journal*, *13*(1), 58–72. http://dx.doi.org/10.2304/eerj.2014.13.1.58

Hernández, F. (2006). El informe PISA: Una oportunidad para replantear el sentido del aprender en la escuela secundaria. *Revista de Educación, Volumen Extra PISA*, 357–379.

INEE. (2017). PISA 2015. In *Programa para la evaluación internacional de los alumnos. Informe español*. Madrid: Instituto Nacional de Evaluación Educativa. Ministerio de Educación, Cultura y Deporte.

Jornet, J. M. (2013). Asignaturas pendientes en las evaluaciones a gran escala. Ponencia en el simposio metodología de investigación en ámbitos educativos plurales y diversos. In *XVI Congreso Nacional y II internacional de Modelos de Investigación Educativa de AIDIPE*.

Koenker, R. (2005). *Quantile regression (econometric society monographs)*. Cambridge: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511754098

Koenker, R. (2017). *Quantreg: Quantile regression. R package version 5.33*. http://dx.doi.org/10.1201/9781315120256

Krapp, A., & Prenzel, M. (2011). Research on interest in science: Theories, methods, and findings. *International Journal of Science Education*, *33*, 27–50. http://dx.doi.org/10.1080/09500693.2010.518645

Lipsitz, M., Belloni, A., Chernozhukov, V., & Fernández-Val, I. (2016). *Quantreg. nonpar: An R package forperforming nonparametric series quantile regression. arXiv preprint arXiv:1610.08329*. http://dx.doi.org/10.32614/RJ-2016-052

Marsh, H., Hau, K., Artelt, C., Baumert, J., & Pechar, J. (2006). OECD's brief self-report measure of educational Psychology's most useful affective constructs: Cross-cultural, psychometric comparisons across 25 countries. *International Journal of Testing*, *6*, 311–360. http://dx.doi.org/10.1207/s15327574ijt0604_1

Mazzeo, J., & von Davier, M. (2013). Linking scales in international large-scale assessments. In L. Rutkowski, M. von Davier, & D. Rutkowski (Eds.), *Handbook of international large-scale assessment: Background, technical issues, and methods of data analysis* (pp. 229–258). New York: Chapman and Hall/CRC.

Newell, A. D., Tharp, B. Z., Vogt, G. L., Moreno, N. P., & Zientek, L. R. (2015). Students' attitudes toward science as predictors of gains on student content knowledge: Benefits of an after-school program. *School Science and Mathematics*, *115*(5), 216–225. http://dx.doi.org/10.1111/ssm.12125

Nugent, G., Barker, B., Welch, G., Grandgenett, N., Wu, C., & Nelson, C. (2015). A model of factors contributing to STEM learning and career orientation. *International Journal of Science Education*, *37*(7), 1067–1088. http://dx.doi.org/10.1080/09500693.2015.1017863

OECD. (2006). *Evolution of student interest in science and technology studies (Policy Report)*. Paris: OECD Publishing.

OECD. (2009). *Equally prepared for life? How 15 year old boys and girls perform in school*. Paris: OECD Publishing.

OECD. (2012). *Are students more engaged when schools offer extracurricular activities? PISA in Focus 18*. Paris: OECD Publishing.

OECD. (2014). *Scaling procedures and construct validation of context questionnaire data*. Paris: OECD Publishing.

OECD. (2016a). *PISA 2015 Results in focus, PISA in Focus 67*. Paris: OECD Publishing.

OECD. (2016b). *PISA 2015 assessment and analytical framework: Science, reading, mathematic and financial literacy, PISA*. Paris: OECD Publishing.

OECD. (2016c). *PISA 2015 Results (Volume I): Excellence and equity in education, PISA*. Paris: OECD Publishing.

Pereira, D., Perales, M. J., & Bakieva, M. (2016). Trends analysis in the investigations realised from the data of the PISA Project PISA. *Relieve*, *22*(1), 1–17. http://dx.doi.org/10.7203/relieve.22.1.8248

Potvin, P., & Hasni, A. (2014). Interest, motivation and attitude towards science and technology at K-12 levels: A systematic review of 12 years of educational research. *Studies in Science Education*, *50*(1), 85–129. http://dx.doi.org/10.1080/03057267.2014.881626

R Core Team. (2017). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Rutkowski, D., & Rutkowski, L. (2013). Measuring socioeconomic background in PISA: One size might not fit all. *Research in Comparative and International Education*, *8*(3), 259–278. http://dx.doi.org/10.2304/rcie.2013.8.3.259

Schleicher, A. (2016). *PISA excellence and equity. PISA day: Global lessons for the U.S. Education system and economy. Alliance for excellent education*. Paris: OECD Publishing.

Schunk, D. H. (1991). Self-efficacy and academic motivation. *Educational Psychologist*, *26*, 207–231. http://dx.doi.org/10.1207/s15326985ep2603&4_2

Schwartz, S. (1994). The fallacy of the ecological fallacy: The potential misuse of a concept and the consequences. *American Journal of Public Health*, *84*, 819–825. http://dx.doi.org/10.2105/AJPH.84.5.819

Stanat, P., Artelt, C., Baumert, J., Klieme, E., Neubrand, M., Prenzel, M., . . . Schneider, W. (2002). *PISA 2000: Overview of the study—Design, method and results*. Berlin: Max Planck Institute for Human Development.

Täht, K., & Must, O. (2013). Comparability of educational achievement and learning attitudes across nations. *Educational Research and Evaluation*, *19*(1), 19–38. http://dx.doi.org/10.1080/13803611.2012.750443

Torney-Purta, J., & Amadeo, J. A. (2013). International large-scale assessments: Challenges in reporting and potentials for secondary analysis. *Research in Comparative and International Education*, *8*(3), 248–258. http://dx.doi.org/10.2304/rcie.2013.8.3.248

Willms, J. D. (2006). *Learning divides: Ten policy questions about the performance and equity of schools and schooling systems*. Montreal: UNESCO Institute for Statistics.