Trends analysis in the investigations realised from the data of the PISA Project

Análisis de tendencias en las investigaciones realizadas a partir de los datos del Proyecto PISA

Pereira, Dora [1]; Perales, María-Jesús [2] & Bakieva, Margarita [2]

(1) Piaget- Institute (Portugal)   (2) Universidad de Valencia (Spain)

Abstract

This study is part of a broader research that aims to design a model of evaluation of educational systems, based in social cohesion. The starting point is located in previous research about the context variables; it is cared more frequently a theoretical justification and metric quality of output variables, leaving aside the variables that can help interpret the results. In our case, we conducted a review of the scientific literature, based in empirical papers that use PISA database and published between 2000 and 2012. The methodology is structured in two phases: identification and selection of documents and further analysis of these documents. The analysis of documents consists of identifying certain indicators in order to explore. Some results suggest that researchers in the field of education underuse PISA databases.

Keywords:

OCDE; PISA; educational systems; empirical research; educational research

Resumen

Este estudio es parte de una investigación más amplia que pretende diseñar un modelo de análisis de evaluaciones de sistemas educativos que permita valorar su aportación para la cohesión social. El punto de partida se sitúa en las investigaciones previas que analizan el papel de las variables de contexto, de entrada y de proceso en las evaluaciones de sistemas educativos, que frecuentemente cuidan mucho más la justificación teórica y la calidad métrica de las variables de producto que estas otras, que finalmente ayudan a interpretar los resultados. En este caso realizamos revisión de los trabajos científico-académicos del corte empírico, publicados entre los años 2000 y 2012, que utilizan bases de datos PISA. La metodología se estructura en dos fases: identificación y selección de los documentos para formar una base y posterior análisis de cada uno de los documentos con el fin de extraer indicadores que nos permitieran realizar un análisis exploratorio. Algunos resultados hacen pensar que las bases de datos PISA están infrautilizadas, sobre todo por los investigadores en el campo de educación.

Palabras clave:

OCDE; PISA; sistemas educativos; investigación empírica; investigación educativa

PISA (Program for International Student Assessment) is widely known by anyone interested in the national and international state of education. As it is known, PISA analyses are used in education in countries which represent 90% of world economy. It serves for monitoring their education systems through the assessment of basic skills at the age of 15, with the aim of providing relevant information to establish policies and practical proposals to reach an effective education (Ozmusul y Atanur, 2013).

PISA is developed in 3 years cycles and assesses linguistic, scientific or mathematical competences, and more recently also evaluates the ability of problem solving and financial competence. Each edition of PISA measures one of three afore mentioned primary domains (Table 1).

Autor de contacto / Corresponding author

Perales, María-Jesús . Dpto. MIDE, Facultad de Filosofía y CC.EE, Av. Blasco Ibañez, 30, 46010-Valencia (Spain). perales@uv.es.

Table 1 - *Areas of PISA by waves*

Waves	Primary Domain	Secondary Domain
2000	Reading Literacy	Mathematical and Scientific
2003	Mathematical Literacy	Reading and Scientific
2006	Scientific Literacy	Reading and Mathematical
2009	Reading Literacy	Mathematical and Scientific
2012	Mathematical Literacy	Reading and Scientific
2015	Scientific Literacy	Reading and Mathematical

The OECD (Organisation for Economic Co-operation and Development) provides the infrastructure, materials and specialists to accomplish the PISA study, although the participating nations or educational participating institutions are required to pay a participation fee. OECD offers PISA databases for free in its official website and any researcher can use them. These databases provide a large amount of information on pupils and schools of the studied countries, the information that extracted with outstanding rigor through a sampling process Besides, these databases offer extensive information on the development indicators, which allow comparing different countries (Cordero, Crespo y Pedraja, 2013).

The OECD offers a lot of reports with explanations of data and demographic inferences, as well as methodological analysis handbooks, and also tutorials about contextual and performance instruments. Indeed, the PISA project enormous evolution is self-evident, and trough the time it has provided much information that initially appears in its reports. So, from 2002 many countries are beginning to generate secondary reports, authored by experts within and outside the OECD, as well as other independent researchers. These reports explain the relationship between the different analysed variables and the results of the assessment. Currently we already have a fairly extensive list of secondary reports, PISA explaining documents, in different languages, that describe the assessment and analysis

methodology, in-depth results, and offer suggestions and recommendations about educational policies based on the analysis of the results.

At the same time, the OECD reports are used as a basis to justify high-impact policies and making decisions about educational policies. Acevedo (2007) considers that international assessments are specially designed for educational managers and administrators – policy oriented studies – in order to enable them to make decisions for the moment and direction of educational reforms. These studies such as PISA, provide empirical evidence to support theoretical and ideological considerations.

PISA tests are designed to assess the pupils' skills and competences in analysing and resolving problems, trying to offer, longitudinally, a profile of the pupils' skills in all countries where it is delivered. Even PISA recognizes its claim to influence education policies, the fact is that also recognizes not to be linked to the curricula, nor to any specific study planes (Gallardo-Gil et al., 2010). It is understood that rather than as a simple artefact of an international benchmarking, it is a complex process of review and reflection for transnational political regulation.

For Carvalho (2009) PISA concerts inquisition and measurement, it is a

dashboard of agreements promotions on what the educational practices and policies are, and which national governments allowed to submit to

external scrutiny, regularly arming the politicians with data and analysis based on models constructed above conventions established by experts (Carvalho, 2009, p. 1017).

Also, there are opinions that suggest that PISA indicators are focussed on comparing the effects and consequences between educational systems, rather than on their results on the dynamic of mutual rendering of accounts. Carvalho (2009) emphasizes that the primary audience of the program are policy makers, so that the information is presented in the reports according to their needs and requests. It's about knowledge for policy making. Thus, it is an origin and final point of political decisions.

PISA produces knowledge about knowledge. The reports carried out to make policy decisions should be the primary and basic material for other secondary analysis conducted by specialists from different areas of knowledge, working out the evidence on student performance in math, science or reading literacy respect for other variables.

Although poorly distributed, and under-consulted and under-understood reports, the most visible results, as the rankings, have a wide social impact and high press tracking (Jornet, 2013; Jornet, 2016; Batista, 2016), being used even to justify reforms in educational legislation.

The critics of PISA reports are numerous. These concern to the technical and methodological aspects of validity of studies (Gorur, 2014), as well as accuracy of measurement of PISA equity (Rutkowski & Rutkowski, 2013) or methodological aspects of comparison between countries (Torney-Purta, 2013), and others (e.g., see Cordero, Crespo y Pedraja, 2013).

Actually, according to Jornet (2013), there are several outstanding questions in large-scale assessment, and among them probably the main question is related to the use of the information that can be obtained from the PISA reports. In the particular case of PISA, the most common criticism is the lack of

explanatory power of assessment studies coupled with the lack of harnessing of the use of assessment databases that the scientific community performs.

This relates to the fact that the number of empirical studies based on data provided form PISA databases is rather limited. Clearly, PISA is not a research project, however, the generated data can be of great interest to researchers (Turner, 2006). Although PISA is a set of large and solid databases, with a great potential for empirical research, few researchers from academic and independent entities tend to use or even consider its use for their research. This apparent lack of commitment of researchers to use and exploit of these data, despite its great social impact, has been the focus of present research, that reviewing the characteristics of the empirical studies that have actually been published with that profile.

Objectives

This paper is a part of a broader study that aims to design an evaluation model of educational systems that allows appreciate their contribution to Social Cohesion. The starting point in this study is located on previous researches about the role of context input and process variables, in the process of educational systems assessment. Those researches consider more the theoretical justification and metric quality of product variables (performance, skills assessment) than other variables that, in the end, will be helpful to explain the results (this branch of study was developed in the project MAVACO with reference EDU2009-13485, funded by the Ministry of Science and Innovation of Spain). The general outlines of the current project, called SECS/EVALNEC are described in Jornet (2012). Briefly, the Assessment Model for Social Cohesion, designed by Jornet, propose a set of instruments with the purpose of interpreting the pupils' outcomes associated to other elements of educational context (De la Orden & Jornet, 2012). Those elements can be, e. g., teaching methodology, the subjective value of education, teacher collegiality and

vision of subjective social justice (Jornet, 2012). This set of elements is composed by indicators of quality of an educational system, based on the concept of Social Cohesion and Inclusivity.

According to the project framework, in this study we conducted an initial revision of empirical papers, using PISA databases and published between 2000 and 2012. These papers should focus on analysing the relationship between the context variables and school outcomes assessed by PISA. The objectives of this research, therefore, have focused on:

- Presentation of conducted literature review process, its criteria and results,
- Presentation of empirical papers that match the criteria for inclusion in the study,
- Character description of the institutions and researchers that conducted studies, according to the affected countries, level of analysis, used statistical methodology and the type of provided recommendations.

Methodology

The methodology of this study is structured into two phases. The first one focused on the creation of documental base of the studying object and identification of essential documents for this study; and the second one is focused on analysis of papers, according to the work protocols, that help to identify quantifiable and/or qualitative data.

Phase 1.

The first phase, as afore mentioned, focuses on the process of identification and selection of the papers to analysis. The methodology of documental revision follows passes provided by Bisquerra (2014), and was focused on identifying the empirical papers that was conducted from academic and research institutions outside the OECD and published between 2000 and 2012, using the PISA databases to research about relationship of context variables and educational outcomes as measured by PISA. The documented base of research is fundamentally constructed with

papers published in scientific journals, as well as papers and/or communications and contributions, working-papers, and e-book chapters, which presents empirical studies using accessible PISA databases for University of Valencia PhD Students[1]. We understand that even if we address systematically search in a wide range of scientific papers databases, surely always there are papers that will not be located, so these papers will not include in the study. In order to minimize this problem, the documentation approach has been structured according to criteria used to select the documentation basis, which were as follows:

- Scope and reputation of the consulted bases;
- Specialization or related to our subject (PISA project reports and/or education systems assessment projects, research about educational organization and educational policy, accountability and educational outcomes, achievement and learning, cross cultural studies, comparative studies, large-scale tests approached educational studies).

Once completed the initial search, a verification search has been conducted from general most popular engines like Google and Schoolar Google, in order to track down other papers that were not indexed in the previously analysed documental bases.

Whenever the documental base was performed to restrict the language, the search was performed using four languages: English, French, Spanish and Portuguese. During the searching process a small set of keywords, in order to identify a biggest number of related papers. The keywords were: PISA, OECD, performance, and context variables.

The purpose was to identify empirical papers, focused on using the PISA databases and published in scientific journals. For this reason, we started by using keyword like PISA and OECD. However, although they are terms

---

[1] All Social Sciences Databases with subscription of University of Valencia, you can see more in www.trobes.uv.es

that are appear very frequently in educational literature, generally, and evaluative, particularly, also is true that most of the found papers refer to the (international o national) reflections about the uses of the results of PISA, its usefulness, social controversy that it provokes, etc. Empirical studies that were conducted using PISA databases were substantially fewer. The keyword number 3 and 4 (performance and context variables) sought to narrow slightly the findings, by elimination the papers focused in theoretical reflection and leaving uniquely empirical papers. For this reason, they cannot be considered keywords that limit searching results too strictly.

The final selection of research papers was based on direct review of the studies, in way that it could verify that the research is in agreement with the aims: to be empirical study using PISA databases. Papers that did not correspond to mentioned type of studies, were discarded. For this reason, it was a very laborious work. We had to address the original sources (scientific papers, communications and working papers, chapters of e-books). However, considering that the quality of research depends largely on the adequacy of the available information, we understood that it was a fundamental work directed to support the validity of this research.

Therefore, there is a bank of documents consisting of 116 works (scientific papers, communications and working papers, chapters of e-books).

In the Table 2 we present different solutions of conducted search in different documental databases. The search was confined between 2000 and 2012.

Table 2 - *Searching parameters*

Source	Boolean searching parameters		
	(PISA) and (OCDE)	(PISA) and (OCDE) and (Contextual variables)	(PISA) and (OCDE) and ((Predictive variables)
Google	59600	5040	12800
Scholar Google	31500	4050	123
troves.uv.es	20	0	0
SpringerLink	32	0	0
Sciencedirect.com	30	0	0
ProQuest.com	942	22	0
persee.fr	12	0	0
Wiley Online Library	11	1	0
Scopus	158	0	0

Phase 2.

Once finished the first phase of the documental search and composed the list of the scientific documents, we proceeded to the next, second phase, focused on the analysis of the information, contained in that scientific papers in order to extract quantitative indicators (data) and qualitative (arguments).

That process allowed us carry out an exploratory analysis based on review of scientific material, using several criteria of expert judgment.

For the descriptive analysis of the information it was used SPSS22 with licence of University of Valencia.

Results

This section describes the results obtained from two phases of study, described previously in the methodology chapter.

Results of Phase 1. Creation of documental database *object of study*: Identification of essential papers for study.

The first phase makes possible identify a large number of papers that published researches using PISA databases. These documents were reviewed to comply with the criteria for inclusion or exclusion in research (ranked by their level of importance):

- Empirical studies,

- Conducted by academic o research institutions, outside of OECD,
- That studied relationship between the variables called "context variables" and performance outcomes measures by PISA, at a time that the performance was considered as variable output.

In the first selection we identified 248 scientific papers. Almost a half of them really satisfied the inclusion criteria to be included in the list of documental review, but many others were excluded for various reasons after detailed analysis. Table 3 shows the different criteria to exclusion and a number of documents.

Table 3 - *Exclusion criteria*

Exclusion Reason	Frecuencia
Not open access document (only abstract)	62
Not an empirical research	21
No use of PISA variables as output variables	13
Another language (not English, Spanish, Portuguese or French)	2
It's official or OCDE report, o national report (from Ministry, Official Statistical Institution)	20
Only on paper format	14
Not a paper/working paper	1

The first criterion for exclusion of found research refers to accessibility. Then, we deleted documents corresponding to official reports of the official institutions like OECD, or its national agencies; because the objective of the study refers to research conducted form academic or research institutions outside the OECD. Finally, the third major reason for exclusion refers to the content of the document: apparently these 248 pre-selected papers referred to empirical research, but a closer scrutiny of the contents revealed that some researches contributed with data, but without empirical analyses, or did not use the PISA results as an output variable.

So, after looking at content review, the documental base was constituted by 116 scientific papers with evidences of scientific quality, based on empirical studies that used PISA databases.

Results of Phase 2. Documental analysis, according to working protocols that allows the identification of quantitative and qualitative information.

Most of the reviewed studies have been published as scientific papers, as we can see in Table 4. Another common format is working papers, along with working papers developed during some academic and scientific or research conferences, and finally the e-book chapter was the less common format.

Table 4 - *Type of paper*

Type	N*	Wave of PISA 2000	2003	2006	2009	Total
Scientific Journal Paper	72	24	21	34	2	81
Working Paper	37	9	15	16	6	46
E-books/ Chapter	7	0	1	5	2	8
Total	116	33	37	55	10	135

Note: Total of analysed research papers of this table (Total, final column) does not equal to the number of works in this study (N, the first column), because the same paper can treat several waves of PISA.

A first description of these papers is focused on the wave of PISA. The PISA Project is organized in successive waves, as it had mentioned before, and those waves explore different areas of knowledge, which allow to complete the entire universe of considered knowledge. Table 5 shows how many of the reviewed documents (presented by year of publication) analyse each of the waves. We can observe a logical period of shortage between the completion of assessments by the OECD and the first independent research. In fact, after four or five years from corresponding PISA waves it begins by a period of maximum scientific productivity in each wave. This is because, firstly, it is takes time to publish in PISA databases, and then it requires time to plan and develop independent research, and, finally, it takes time to publish.

Table 5 - *Year of publication by wave and total scientific papers*

	Total (n=116) F	%	2000 F	%	2003 F	%	2006 F	%	2009 F	%
2001	2	1,72	1	3,03					1	10,00
2003	2	1,72	2	6,06						
2004	3	2,59	3	9,09						
2005	7	6,03	6	18,18	1	2,70				
2006	5	4,31	2	6,06	4	10,81				
2007	11	9,48	6	18,18	6	16,22				
2008	10	8,62	3	9,09	7	18,92	2	3,64		
2009	10	8,62	2	6,06	5	13,50	7	12,73		
2010	28	24,14	5	15,15	7	18,92	21	38,18		
2011	28	24,14	2	6,06	7	18,92	17	30,91	6	60,00
2012	10	8,62	1	3,03			8	14,55	3	30,00
Total	116	100,00	33	100,00	37	100,00	55	100,00	10	

Note: Percentages are calculated per column, being 100% the total of scientific papers for each wave of PISA and column of Total is the sum of all the scientific papers of the four waves of PISA. The same scientific paper can treat two or more waves of PISA, so the sum per row would not be equal to the total works published in that year.

When we focus on the affiliation of the researchers, we can observe that there are a lot of them from working collaboratively from several institutions, which explains why the total number of countries of formant institutions largely exceed in number of 116. In addition, it is evident that there is a large preponderance of European researchers, as describes Table 6.

Table 6 - *Researchers institutions location*

Region	Total	2000	2003	2006	2009
Europe	113	32	39	49	10
Australia and New Zealand	6	2	0	2	2
Asia	14	5	5	8	0
North America	13	5	1	6	1
Central and South America	2	0	1	1	0
Ibero-America	34	2	10	20	6
International organizations	1	0	1	0	0

Note: The boxes contain frequencies of countries where research institutions that provide cover for researchers (authors) are located. These countries are grouped by regions by geographic, linguistic, and economic factor. Thus, the regions of Europe and Ibero-America are grouping countries of Europe, Latin America with Spanish and Portuguese official language. In some cases, the total number of the row exceeds the sum for each wave of PISA because some studies are conducted in collaboration between institutions situated in different countries. The "International organization" category refers to the World Bank.

The Table 6 clearly shows that in most of the studies are involved researchers from European institutions (113 in total). This may indicate a higher sensitivity in Europe about international studies such as PISA (against the US, e. g., which possibly will keep as the most significant regarding their national studies), and also a higher tendency and/or facilities to develop collaborative works involving entities from different countries.

The countries in Central and South America are represented only by institutions of two countries, Chile (Centre for Advanced Research in Education, CIAE) and Colombia (School of Finance and Administration and Institute of Technology).

Regarding this under-representation, Martinez Rizo (2006, p. 154) explains that "like the majority of developing countries, Latin American countries have not an important tradition in this field". This author also underlines that the importance of the participation of these countries must be "as active as it's possible" (p. 166).

In the 116 selected papers there are researchers from 34 institutions of Latin America region. But, as mentioned, only 2 of these are located in Central and South America, being 28 from Spanish and other 4 Portuguese institutions.

Finally, North America hosts 13 research papers in total, followed by Orient countries with 14 research, and 6 from Australia and New Zealand.

The publication language might seem an exclusion criterion of research carried out in regions of Asia However, by including the four most common languages in the world (English, French, Portuguese and Spanish), it has been raised as an inclusive principle. However, the selection of consulted databases itself can act as exclusionary criterion, then they may collect deficiently the research of countries outside of Europe and North America.

In most of the regions, and in the total data, the number of research increases with successive waves of PISA (being especially evident in Europe and Latin America). This may be a reflection of the growing interest towards the PISA Project, in view of its political and social impact, and also greater actual accessibility of databases. However, in some less frequently regions this trend is broken in case of PISA 2003 and it is difficult to find a first explanation for this fact.

Table 7 shows the count of the areas of journals in which the research papers was published. It can be observed a clear tendency to the journals with economic cut, which conditioning the researches' character and the scope of research publication.

Table 7 - *Scope of journal*

Year of publish.	Education	Economy	SS.CC and Psychology*	Engineering	Sociology	Total
2001		1	1			2
2003		2				2
2004		3				3
2005	4	2	1			7
2006	3	2				5
2007	1	8	2			11
2008	2	6		1	1	10
2009	3	6	1			10
2010	12	13	3			28
2011	8	17	3			28
2012	3	4	3			10
Todos	36	64	14	1	1	116

Note: *Includes Social and Behavioural Sciences

Although the OECD is clearly an economic organism, the PISA Project is an educational study. Nevertheless, Table 7 shows that more than 50% of analysed works were published in journals or publications with Economic scope; and Education, was the second cluster, then a 25% remained. This may show that educational researchers react late and with low impact, at least in terms of number of publications. There are few researchers in educational scope which engaged into using PISA databases in their research.

Researchers introduced, now is relevant to the analysis of different regions studied in various works. Thus, in Table 9 we can observe the classification of scientific works based on the studied region. The data clearly show that the most of the works focus on European countries.

Table 9 - *Region focus of the study*

Región	Total (n=116)	2000 (n=33)	2003 (n=37)	2006 (n=55)	2009 (n=10)
Central and South America	45	19	10	18	3
North America	49	19	16	19	2
Australia and Oceania	61	20	20	23	2
Asia	104	24	41	50	5
Europe	733	239	226	301	31
Ibero-America	127	39	35	57	9

Note: n – number of relevant scientific studies by the PISA wave by the Total (first column). The boxes contain the number of focus countries in each case (by Total o by corresponding PISA wave). The sum per row does not match the value in the Total column because some reviewed studies treat multiples PISA waves. The sum by columns does not coincide with the Total N (by wave and Total) because many studies include multiples focus countries in their studies. Ibero-America countries grouping Europe and Latin America countries with Spanish and Portuguese official language.

As it can be seen in Table 9, the studies focus basically on Europe, being Spain, Germany, Italy, Finland, and the UK the most frequently mentioned as a central element of studies. Any research that includes some more recent publications probably may show a greater preference for countries systematically positive-assessed by PISA in the international rankings, such as Finland. In the analysed period, it is presumable that the priority is given to analysis criteria like origin country of researchers or comparing with significant regions.

Ibero-America is by far the next most studied region with 127 items listed as focus countries in 116 papers. However, an analysis confirms that Spain, and with lesser degree Portugal, have a large presence, so the frequency of mentions of focus countries in Central and South America is the 45. Among them, Mexico is most mentioned as focus country, with 25 occasions; and it is consistent with the participation of Mexico in PISA Project, along with Brazil, since its first edition. This confirms the importance for these countries to participate in international studies such as PISA, which makes visible their education systems in line raised before by Martinez Rizo (2006).

In Asian regions the countries such as South Korea and Japan are the most frequently analysed, also it is highlights the lack of China, which chosen not to participate in PISA (except specific regions of China, such as Hong Kong and Macao).

Regarding the following question about reviewing the content of conducted research two descriptive key-elements have been considered. First, the level of analysis, and then the type of analysis used during the research.

Table 10 provides data about the level of analysis of empirical research. PISA Reports, and the databases that they offer, are structured according to different levels of analysis, allowing different types of data grouping, which are linked to different research goals. Although the study does not allow make conclusions in terms of pupils or specific schools, because it is not designed for it, but the bases provides information about pupils grouped by centres. This will allow analysis associated with different individual collected variables (such as demographic, social, economic characteristics, school trajectory and others), or functional characteristics of schools (such as ownership or size).

The following levels of analysis are the country and the region (considered as a geographical criterion for countries grouping).

At first sight it can be seen that most of the scientific papers develops research at the level of pupil or group of pupils, followed by number of research papers that studying variables at level of schools or type of schools, followed that by variables at level of countries or communities (grouping of several local counties in one country), and finally research at level of region (grouping of several countries).

Table 10 - *Level of analysis*

Level of analysis	Total (n=116)		2000 (n=33)		2003 (n=37)		2006 (n=55)		2009 (n=10)	
	F	%	F	%	F	%	F	%	F	%
Pupil/s	115	99,14	0	0,00	0	0,00	54	98,18	0	0,00
School	99	85,34	28	84,85	35	94,59	44	80,00	9	90,00
Country	34	29,31	7	21,21	12	32,43	17	30,91	5	50,00
Region	27	23,28	10	30,30	10	27,03	11	20,00	1	10,00

Note: Level of analysis: Pupil – by groups of pupils, School – by types of schools or educational centres, Country – by several counties of same country, Region – by several countries in the same geographic o linguistic region. The percentage is regard to the total number of scientific works of each PISA wave and total of the all analysed scientific papers.

In Table 10 we can see that the tendency of types of analysis is different, depending on wave of PISA. In works that use data of PISA 2000, 2003 or 2009 there are no studies contrasting variables at level of pupil or group of pupils. However, studies based on data of PISA 2006 presents the 98,18% of cases with this type of variables or indicators.

Obviously, the different levels of analysis are not mutually exclusive of each other. In this regard, we can observe that with the advancement of time the analyses of successive waves combine different levels of analysis. Thus, while the analysis of PISA 2000 is performed primarily at school level (and the low duplication of frequencies indicates that only are made at that level), in the analysis of the PISA 2006 we can see that it is carried out simultaneously at the level of pupil and school, combining information.

For the description of statistical analysis used in each case we raised five categories, which are not mutually exclusive of each other:

- Univariate and bivariate descriptive analysis, providing a comprehensive presentation of the studied variables,
- Hypotheses contrast to verify differences or similarities between established groups with different characteristics, and analyse the influence of some other variables. Among the contrast of hypothesis research, we verified whether or not they value the effect size. This is considered especially relevant as a safeguard for the research, because the large size of the samples of PISA databases can influence the results of contrast of hypothesis.

- Multivariate analysis of different types. Group together multiple variables to study the combined influence among them. Depending on research, it can be Multivariate analysis of variance, Multiple Regression Analysis, Factor Analysis, Discriminant Analysis, Cluster Analysis, Multi-dimensional Scaling and/or Causal Models.
- Multilevel analysis. They form part of a group of statistical analysis models that allow treat nested data sets within a population with hierarchical structure, understanding that the different levels of hierarchy are levels of analysis (Amador and Lopez-Gonzalez, 2007). Figure 1 shows an example of structured data on two levels.
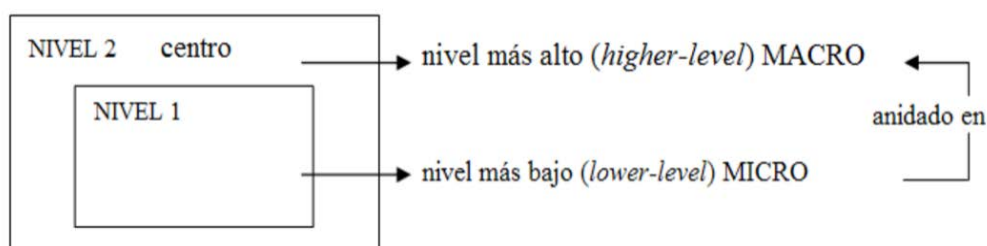

Figure 1. Structure of the data on multilevel analysis.
Source: Amador & López-González (2007)

Table 11 provides aggregated data on the type of analysis used in research. *Grosso modo*, the advanced level classification of statistical procedures not become advanced in most cases. In fact, in line with our study, Pey-Yan and Yi-Chen (2015) found that the effects of size and design, essential adjustments to analyse data from large-scale studies, are detected in less than half of cases. Nevertheless, tests of statistical significance may be insufficient in studies of this magnitude, were the observed size effect is fundamental. Statistical procedures of size effect are designed in order to quantify that effect and they are needed in such extensive population studies (Coe and Merino, 2003). Its

relative absence in these analyses is relevant because it may mean that they can question some of the results, and it can be indicative of

lack of tradition in the analysis of large samples made by entities outside the important international organizations.

Table 11 - *Type of analysis used in the study*

	Total (n=116)		2000 (n=33)		2003 (n=37)		2006 (n=55)		2009 (n=10)	
	F	%	F	%	F	%	F	%	F	%
Descriptive	65	56,03	17	51,52	25	67,57	31	56,36	6	60,00
Contrast of Hypothesis	80	68,97	27	81,82	30	81,08	35	63,64	4	40,00
-- Size effect	38	32,76	12	36,36	17	45,95	12	21,82	3	30,00
Multivariate	82	70,69	23	69,70	26	70,27	35	63,64	9	90,00
Multilevel	62	53,45	17	51,52	17	45,95	26	47,27	7	70,00

Note: The percentage is regarding the total number of papers by wave of PISA and total number of analysed works. Some studies use two or more levels of analysis during the research.

Multivariate analyses are the most common in the revised research (more than 70% of papers use them, being clearly more frequent in the wave of PISA 2009). They allow a very relevant approach in the educational phenomenon, which is very difficult to define. Multivariate analysis allows reflecting on the complexity of education in appropriate way, because it treats the influence and the interdependence of different variables globally from different approaches.

One of the most important consequences of PISA Studies are the recommendations and reflections that can be drawn. Therefore, the last category of analysis of the 116 studied papers will be it. Table 12 shows the trend of kinds of recommendations, by the year of publication. Clearly, the number of recommendations is increasing over the years, showing quite a high growth.

Table 12 - *Type of recommendations*

Year of publication	Total*		Recomm. About Educational Policy		Recomm. About Curricula		Recom. About PISA Methodology		Recomm. For Future Researchers	
	F	%	F	%	F	%	F	%	F	%
2001	2	1,72	1	0,86	0		1	0,86	1	0,86
2003	2	1,72	2	1,72	0		0		2	1,72
2004	3	2,59	2	1,72	2	1,72	2	1,72	2	1,72
2005	7	6,03	6	5,17	2	1,72	2	1,72	2	1,72
2006	5	4,31	5	4,31	3	2,59	1	0,86	2	1,72
2007	11	9,48	10	8,62	3	2,59	4	3,45	5	4,31
2008	10	8,62	7	6,03	6	5,17	6	5,17	4	3,45
2009	10	8,62	5	4,31	4	3,45	2	1,72	2	1,72
2010	28	24,14	25	21,55	10	8,62	11	9,48	18	15,52
2011	28	24,14	25	21,55	12	10,34	2	1,72	13	11,21
2012	10	8,62	6	5,17	2	1,72	2	1,72	6	5,17
All	116	100,00	94	81,03	44	37,93	33	28,45	57	49,14

Note: The last column contains the total number of research papers per year. It need not be equal to the sum of the corresponding row, because in the same paper may occur recommendations of different types. The percentage is calculated form the total number of works (n=116).

The following of the first application of PISA year there are two studies are published. In them, we found recommendations on educational policy, PISA methodology (assessment or data analysis), and also recommendations for future research. About 81% revised documents provide recommendations on educational policy, being much lower percentage of works that provide review about teaching strategies, while we understand that it is a mucho more specific area. Yet it is significant to note that although only 25% of studies have been published in the journals with scope on education, researchers have found sufficient findings to make recommendations about the area that initially not theirs. This interdisciplinary potential of educational systems assessment studies should be supported and considered, in order to be a really constructive way to treat the phenomenon of education; but at the same time the voice of educational experts cannot be overlooked.

It is also important to note that almost 50% of the researches provide recommendations for future research. Often,

finish of research is the right time to raise another, therefore the questions that arise from the obtained results, are methodological analysis-based guidelines that have been able to verify or disprove. In this sense, it is essential to emphasize the cumulative nature of the research lines and the need to promote research networks to deepen collaboratively on the most relevant topics.

Finally, also highlights that almost 30% of reviewed studies make reflections about the PISA methodology. The technical difficulty of PISA Study and its large political and social impact (which also has been growing through the time) make necessary a constant review of its own methodology, involved in the different phases (from identification of the construct and universe for measure, to analysis and dissemination of results). Also it needs be powered by thoughts and ideas form researchers of PISA team, but also from external researchers. In this sense, these methodological reflections should be considered by the PISA team as a treasure to be thought carefully.

Table 13 - *Type of recommendations*

| | Type of recommendations | | | | | | | | | |
| | Educational Policy | | Curricula | | PISA Methodology | | Future Research | | Total (100%) |
Scope	N	%	N	%	N	%	N	%	
Education	32	88,89	22	61,11	11	30,56	22	61,11	36
Finance	54	84,38	17	26,56	15	23,44	28	43,75	64
SS.CC.*	8	57,14	5	29,41	6	42,86	5	29,41	14
Engineering					1	100,00	1	100,00	1
Sociology							1	100,00	1

Note: SS.CC. also includes Psychological and Behavioral Sciences

As for the different types of recommendations made by researchers in each analyzed study, in the Table 13 we can observe how they are distributed in terms of scope of the journal were it was published.

There is only one journal with Sociological scope, and another one in field of Engineering, and both of publications focus their recommendations in comments about the future research, or questions about the PISA

methodology. If we go to the most frequent publications with scope on Economy and Education, the situation changes. Comparatively, the documents which wwere published in journals with scope on Education are most prolific in all kinds of recommendations. Recommendations concerning education policy are the most frequent, being present in almost 90% of published studies in the field of Education,

but interestingly enough, also in more than 80% of the works situated in the Economic sphere. The curricula recommendations are more in the studies published in the sphere of Education (61,11% compared to 26,56% of the studies published in Economic sphere). Certainly, large educational policies are very interdisciplinary, and the economic sphere has its own perspective on this. The recommendations about the way of giving classes, teaching, the classroom environment, are much more frequent in education journals, although when the PISA may seem a remote to classroom study. Both type of publications, scoped on economy and education, offers less frequently recommendations on the methodology of PISA, but they are present in about 25% of studies. The complexity of PISA makes difficult to make contributions about its methodology, but these may be especially relevant when formulated from the position of independence and externality of researchers, specific areas of interest and particular questions, and the expertize on educational systems assessment.

Conclusions

The presented study considers the awareness of the quality and exhaustiveness of PISA databases, and the feeling that they are underused. And the results come to confirm this suspicion, outlining some possible causes, with the aim to offer routes to boost their use.

If, according to Carvalho (2009), the main virtue of PISA is to produce knowledge for policy, this can be considered important especially from outside of OECD institutions, which might be interested in the research about educational questions, and also collaborate to this process from other perspectives and priorities. That is why the opportunity that provides the PISA databases should not be neglected.

The OECD has improved the availability of data. The databases are becoming easier to access, and there are tutorials of PISA with instructions on how to perform analysis properly. Nevertheless, the bases remain complex, and data long-winded, as in number of constructs and variables, as in number of subjects. This is probably the main drawback of the handling of the bases. Its biggest advantage also its biggest difficulty.

The study whose findings we present is based on a documental search process. Available documents were sought, using English, French, Portuguese and Spanish languages. Also documents must show the results of empirical studies developed from PISA databases, where the PISA performance were product variables, or analyzed in terms of other context variables. The presented conclusions are based on the quality of this search. Similarly, it can show that the consulted documental bases probably represent better European and American publications than others such as Asian or Oriental scientific publications, which are probably under-represented in a biased way. Limitations of the study, therefore, are linked to the used keywords, as well as consulted bases and availability of studies, since the point of departure is the institutional scientific documental databases (University of Valencia).

The review shows that the analyzed studies are conducted mainly from institutions with European linkage, and also focused on European zone. The published works especially refers to the economic sphere, and considerably lower the educational area, after a usual grace period of five years from the databases publication.

Europe's preponderance as a researcher and researched context reveals a particular interest in this region for this type of study. PISA is contextually relevant for Europe, because of study that provides interesting information for their own education policies in a context familiarized to the collaboration and agreement to design and joint policies. In the same way, the European research institutions have the resources and people with formal training and tools to analyze such databases, and also used to collaborative dynamics of research. In other regions of the world, these situations do not occur in the

same way and, for a lower interest and/or reduces availability of possibilities, are less involved in conduct and publish empirical studies of this type.

Regarding the used methodology of analysis, contrast of hypothesis and multivariate studies are the most common. Progressively, methodological strategies that further respect the complexity of the educational phenomenon are being implemented. However, it seems necessary a methodological reflection on the used strategies and the way to do this. Thus, the evidence that only one third part of the studies which were conducted contrast with the hypothesis that control the size effect (being something fundamental as large as the PISA samples), confirm the need to reflect on how to carry out the studies. PISA's own team, attentive to studies that publish in this regard, has already published meanwhile instructions on how to perform this kind of empirical studies.

In this same line, the relatively low presence of educational studies can be linked to a lack of training of professionals and researchers in educational methodologies to approach this type of analysis. The weight of the local and the contextual information is essential in education, and cannot hide the necessity on researchers with sensitivity, interest and educational training to tackle and to engage in major international assessments, and can make analysis from its data.

All reviewed documents make recommendations of different kinds. Most part of the recommendations focuses on policy measures, followed by recommendations for future research and then teaching and curricula recommendations. To a less degree these recommendations focus on issues related to the methodology of the PISA Study. This may be because of the great difficulty of interpreting methodology used in the PISA Study, given the complexity of developing questionnaires and methodology analysis used, which in turn strengthens the need for methodological specialized training to carry out this type of analysis.

From this perspective, the involvement of OECD outside researchers' teams, sometimes multidisciplinary, often from different institutions and countries. These teams can help deepen all available information, providing specific studies on specific variables in which each team can be a specialist, or providing comparative studies, not really addressed by PISA, and allowing relevant conclusions on particular topics.

To do this, efforts from OECD for the clarity and availability of data should be a must. Despite the progress made in this regard, the bases are underused for empirical studies, and not always used appropriately. Additional measures form PISA organization can help remedy these shortcomings. On the other hand, from the research institutions the relevance of their researchers to deeper into this data and provide ways of deepening and analysis must be demonstrated. Excessive simplicity in the PISA findings can be frequently found, and the big social impact of these simplistic readings, should be a crucial incentive for empirical based research contributions to make a proper reading of this study, making the readings counterweight, interested and sometimes biased. Thus, the final joint objective, should aim at gaining a better representation of what happens in educational systems and how they can contribute collaboratively to make better education for everyone.

References

Amador, M. & López-González, E. (2007). Una aproximación bibliométrica a los modelos multinivel. *RELIEVE, 13*(1). Retrieved from: http://www.uv.es/RELIEVE/v13n1/RELIEVE v13n1_3.htm

Batista, J. (2016, enero 26). Cómo desmontar el informe PISA. *Las Provincias.* Retrieved from: http://www.lasprovincias.es/comunitat/201601/26/como-desmontar-informe-pisa-20160126001834-v.html

Bisquerra, R. (Coord.). (2004). *Metodología de Investigación Educativa*. Madrid: Muralla.

Bottani, N. (2006). La más bella del reino: el mundo de la educación en alerta con la llegada de un príncipe encantador. *Revista de Educación*, extraordinário, 75-90.

Carvalho, L.M. (2009). Governando a Educaçao pelo espelho do perito: Uma análise do PISA como instrumento de regulaçao. *Educ. Soc., Campinas, 30*(109), 1009-1036.

Coe, R. & Merino, C. (2003). Magnitud del Efecto: Una guía para investigadores y usuarios. *Revista de psicología de la PUCP, 21*(1), 146-177.

Cordero, J., Crespo, E. & Pedraja, F. (2013). Rendimiento educativo y determinantes según PISA: Una revisión de la literatura en España. *Revista de Educación*, 362, 273-297. doi: http://dx.doi.org/10.4438/1988-592X-RE-2011-362-161

De la Orden, A. & Jornet, J. (2012) La utilidad de las evaluaciones de sistemas educativos: el valor de la consideración del contexto. *Bordón, 64*(2), 69-88.

Gorur, R. (2014). Towards a Sociology of Measurement in Education Policy. *European Educational Research Journal, 13*(1), 58-72. doi: http://dx.doi.org/10.2304/eerj.2014.13.1.58

Gallardo-Gil, M., Fernández-Navas, M., Sepúlveda-Ruiz, M.P., Serván, M.J., Yus, R. & Barquín, J. (2010). PISA and scientific competence: An analysis of the PISA tests in the Area of Science. *RELIEVE, 16*(2). doi: http://dx.doi.org/10.7203/relieve.26.2.4138

Jornet, J. (2012). Dimensiones docentes y cohesión social: reflexiones desde la evaluación. *Revista Iberoamericana de Evaluación Educativa, 5*(1), 349-362.

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

Jornet, J. (2016, enero 26). La educación no está tan mal; el informe PISA está sobrevalorado. *Levante, el Mercantil Valenciano.* Retrieved from: http://www.levante-emv.com/comunitat-valenciana/2016/01/26/educacion-mal-informe-pisa-sobrevalorado/1370614.html

Martínez Rizo, F. (2006). PISA en América Latina: lecciones a partir de la experiencia de México de 2000 a 2006. *Revista de Educación, Número extraordinario*, 153-166, Retrieved from http://www.oei.es/evaluacioneducativa/pisa_america_latina_martinez_rizo.pdf

Ozmusul, M. & Atanur, G. (2013). School Policies and Practices at Upper Secondary Schools in Turkey. According to PISA 2009 Data. *Mevlana International Journal of Education, 3*(4), 186-199. doi: http://dx.doi.org/10.13054/mije.13.69.3.4

Pey-Yan, L. & Yi-Chen, H. (2015). Statistical techniques utilized in analyzing PISA and TIMSS data in Science Education from 1996 to 2013: A methodological review. *International Journal of Science and Mathematics Education,* 13, 1449-1468. doi: http://dx.doi.org/10.1007/s10763-014-9558-5

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. doi: http://dx.doi.org/10.2304/rcie.2013.8.3.259

Torney-Purta, J. (2013). International Large-Scale Assesments: challenges in reporting and potentials for secondary analysis. *Research in Comparative and International Education, 8*(3), 248-258. doi: http://dx.doi.org/10.2304/rcie.2013.8.3.248

Turner, R. (2006). El Programa Internacional para la Evaluación de los Alumnos (PISA). Una perspectiva general. *Revista de Educación, Número extraordinario,* 45-74. Retrieved from: http://www.revistaeducacion.mec.es/re2006/re2006.pdf .

NOTES

| Authors / Autores	To know more / Saber más
Pereira, Dora (doralousadopereira@gmail.com). Professor and researcher at the Research in Education and Community Intervention (RECI) (Piaget- Institute -Portugal). Specialist in Educational Assessment. Author of several articles in the field of evaluation of education systems, psychoeducation and gender. Member of the Group of Medición y Evaluación (GEM- UV), Member of RIIED and RIMESE	Google académico ResearchGate
Perales, María-Jesús (perales@uv.es). Pedagoga, Doctora en Ciencias de la Educación. Profesora Titular del Departamento de Métodos de Investigación en Educación de la Universidad de Valencia (España). Miembro del Grupo de Evaluación y Medición (GEM-UV), sus principales líneas de investigación son el desarrollo de planes de evaluación en ámbitos sociales y de cooperación internacional, y el desarrollo de pruebas de medición para variables socioeducativas. Su dirección postal es: Dpto. MIDE, Avda. Blasco Ibáñez, 30. 46010-Valencia (España)	Google académico
Bakieva, Margarita (margarita.bakieva@uv.es). Máster en Política, Gestión y Dirección de los Centros Educativos y Licenciada en Pedagogía por la Universitat de València. Doctoranda en Educación en la Universidad de Valencia. Miembro del Grupo de Evaluación y Medición (GEM- UV). Desarrolla su investigación en torno al tema de evaluación de la dimensión de la Colegialidad Docente, en el marco de la evaluación de Sistemas Educativos para la Cohesión Social. Su dirección postal es: Dpto. MIDE, Avda. Blasco Ibáñez, 30. 46010-Valencia (España)	Google académico ResearchGate academia.edu
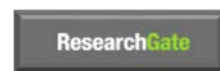
RELIEVE

Revista **EL**ectrónica de **I**nvestigación y **EV**aluación **E**ducativa
E-Journal of Educational Research, Assessment and Evaluation
[ISSN: 1134-4032]