

Gallardo-Gil, Monsalud; Fernández-Navas, Manuel; Sepúlveda-Ruiz, María-Pilar; Serván, María-José; Yus, Rafael & Barquín, Javier (2010). PISA and scientific competence: an analysis of the PISA tests in the Area of Science. *RELIEVE*, v. 16, n. 2, p. 1-17. http://www.uv.es/RELIEVE/v16n2/RELIEVEv16n2_6eng.htm

**e-Journal of Educational
Research, Assessment and
Evaluation**



**Revista ELectrónica de
Investigación y EValuación
Educativa**

PISA AND SCIENTIFIC COMPETENCE: AN ANALYSIS OF THE PISA TESTS IN THE AREA OF SCIENCE

*[PISA y la competencia científica: un análisis de las pruebas de
PISA en el área de ciencias]*

by/por

[Article record](#)

[About authors](#)

[HTML format](#)

Gallardo-Gil, Monsalud (mongalgil@uma.es)

Fernández-Navas, Manuel (mfernandez1@uma.es)

Sepúlveda-Ruiz, María-Pilar (mdsepulveda@uma.es)

Serván, María-José (servan@uma.es)

Yus, Rafael

Barquín, Javier (barquin@uma.es)

[Ficha del artículo](#)

[Sobre los autores](#)

[Formato HTML](#)

Abstract

This work is part of an investigation focusing on the educational evaluation of the skill of Learning how to Learn, which addresses the issue of basic skills that students need to develop to function throughout life in the contemporary social context. One purpose of the study is to analyze the potential of external and diagnostics evaluations to estimate the level of acquisition of these skills. Focusing on a methodology of qualitative and quantitative analysis of the PISA tests released in the specific area of Science (2000-2006), we conclude that these tests require scientific skills of low complexity, with marked presence of mere reproduction. In addition, we appreciate the gap between the concept of competence proposed DeSeCo and what PISA really evaluates according to the results obtained in the area of Science.

Keywords

Assessment; Evaluation; Education; Skills; Competence; PISA; Sciences.

Resumen

El presente trabajo forma parte de una investigación centrada en la evaluación educativa de la competencia de Aprender cómo Aprender, en la que se aborda la temática de las competencias básicas que los estudiantes necesitan desarrollar para desenvolverse durante toda la vida en el contexto social contemporáneo. Uno de los propósitos del estudio es analizar el potencial de las evaluaciones externas y diagnósticas para estimar el grado de adquisición de estas competencias. Centrándonos en una metodología de análisis cualitativo y cuantitativo de las pruebas liberadas de PISA en el área específica de Ciencias (2000-2006), concluimos que estas pruebas demandan capacidades científicas de baja complejidad, con acentuada presencia de la mera reproducción. Además, valoramos la distancia existente entre el concepto de competencia propuesto en DeSeCo y lo que realmente parece evaluar PISA a la luz de los resultados obtenidos en el área de Ciencias.

Descriptores

Evaluación; Educación; Competencias; PISA; Ciencias..

Introduction

This study is part of the research being carried out by the HUM-311 group of the School Organisation and Didactics department of the Universidad de Málaga on the educational assessment of the skill of Learning how to Learn [1]. This project addresses the issue of basic skills as a potential mechanism for the development of life skills, and particularly higher order skills which can be grouped under "Learning how to Learn" (Hargreaves, 2005). One of the purposes of this study is to analyze the potential of external and diagnostic evaluations to estimate the level of acquisition of these competencies and skills. These evaluations include inter alia, the OECD Programme for International Student Assessment (PISA), which bases its foundation on the holistic concept of basic competencies defined in the DeSeCo Report for the OECD (2002), which also served for the establishment of The Recommendation of the European Parliament and the Council of 18 December 2006 on key competencies for lifelong learning. This Recommendation, in turn, was included in the new Spanish Act on educational reform: Organic Act 2/2006 on Education –LOE– (Ministry of Education, 2006). Based on this framework, our aim is to study and assess the PISA Project tests of basic competencies. For this purpose, specific working groups were created within the research team for each area of the curriculum. This article summarizes the main conclusions in the area of Science, i.e. the so-called scientific competence referred to in the Spanish Organic Act on Education: "*Competence in knowledge of and interaction with the physical world*".

Although the PISA assessment project has shifted the foundations of our educational system since its first report in the year 2000, it was particularly in 2006 (OECD, 2006), the year in which the report was dedicated to the assessment of Science, that it began to have a remarkable impact on this area of the curriculum. However, until now there seems to have been more concern for explaining the nature and foundation of the tests (e.g. Harlem, 2002; Oñorbe De, 2008) and the reading

of the results and their implications (Camano, 2007; Puente Azcutia, 2008). The main lines of research relate not only to the results obtained, but also to the evaluation technique followed. For example, a critical aspect of the latter lies in the difficulty of assessing scientific knowledge through paper and pencil tests (Psalidas et al., 2008). Other studies have focused on analyzing curricular materials, such as Hatzinikita et al. (2008), who found that there was a large gap between the type of tasks proposed by PISA and the textbooks used in educational centres, which in part would explain the low results obtained on these tests. In a diametrically opposing line, there is a long list of researchers who have focused on the PISA tests as an opportunity to improve current teaching practices. For example, Hernández (2001) saw the PISA reports as a perfect opportunity for rethinking the tasks carried out in schools, evidencing that schools tend to teach for reproduction purposes, while the PISA tests are aimed at the application or transfer of knowledge to everyday situations. In a similar line, but specifically in the area of Science, Gil and Vilchez (2001) have held in various forums that the PISA reports as a whole are a potentially valuable instrument for improving learning, teaching and the curriculum. Other research has focused on current problems relating to the adjustment of the curriculum in our educational system to the competencies that PISA aims to assess, concluding that the curriculum of the LOE has not been developed to address teaching of the type that underlies all the competencies assessed by PISA, there being mainly procedural areas of science education that have not been referred to under Law.

Taking into account this scenario, it is clear that there is hardly any research that has analytically approached the contents of the PISA tests in an attempt to thoroughly analyze whether the scientific competencies are truly assessed in the holistic sense proposed by the DeSeCo report of the OECD. This is the aim of the work presented in this article, which assesses whether the PISA tests can be considered as useful tools for assessing skills and competencies within the general strategies of the construct of "Learning how to Learn".

PISA in the area of science

As is well known, the PISA comparative assessment model is characterized as being a sample (15 year-old students) and cyclical (triennial) survey comprising three main assessment areas: reading (2000), mathematics (2003) and science (2006). Under this model, the test is designed to identify students' abilities to analyze and solve problems, so as to offer a profile of the performance of fifteen year-old students from all the countries where applied. Additionally, although PISA acknowledges that its fundamental intention is to affect educational policies, it must be borne in mind that it also acknowledges that it is not linked to general or specific curricula. PISA's aim is ultimately to assess 'competencies', which is done empirically in the area of science by questioning students on their ability to identify scientific questions, to explain scientific phenomena and to use scientific evidence. In this sense, within the theoretical foundations of PISA in the area of science, noteworthy is the fact that the assessment approach focuses on the 'application' of knowledge versus the 'rote memorization of concepts' (OECD, 2006).

It should be noted that in the area of Science, PISA's theoretical foundations have apparently evolved since 2000, in which it referred to "scientific education" without making any reference to "competencies (OECD, 2002). Subsequently, following the DeSeCo report (2002), in PISA- 2003, the notion of "scientific competence" began to be addressed, although in the same terms as in 2000, i.e. as a summation of knowledge, processes and situations or contexts (personal, public and global) (OECD, 2004). Finally, in 2006, the specific year in which the assessment focused on the area of science, the concept of "scientific competence" applied to a specific individual was introduced, while maintaining the objective to assess 'knowledge' (concepts) and its application to a situation or context (skills), the only new addition being students willingness or attitude towards the tests and scientific knowledge (OCDE, 2006).

PISA-2006 attempts to demonstrate scientific competencies through the domain of the scientific procedures which form the basis of the questions, the understanding of skills required to solve them and the evaluation of the students' attitudes towards science today. Furthermore, PISA assesses scientific knowledge in three dimensions: (a) scientific processes or skills; b) scientific concepts and contents; and c) the context in which scientific knowledge is applied. Additionally, PISA identifies five scientific processes: recognizing issues that it is possible to investigate scientifically; identifying the evidence which is necessary in a scientific investigation; drawing or assessing conclusions; and communicating valid conclusions which evidence an understanding of scientific concepts in certain situations. These scientific processes are organized into three groups of competencies according to the predominant type of thinking skills required to answer the questions posed: a) Description, explanation and prediction of scientific phenomena; (b) Understanding of a scientific investigation and c) Interpretation of scientific evidence and conclusions. This compartmentalization of the competency into the areas and skills which PISA finally intends to assess implies, in our view, a distancing from the holistic concept of competencies as defined by DeSeCo, it being difficult to conclude that this analytical system evaluates scientific competence in the terms defined by PISA: *"The capacity to use scientific knowledge, to identify questions and to draw evidence-based conclusions in order to understand and help make decisions about the natural world and the changes made to it through human activity"* (OECD, 2006).

Description of the methodological process

As already discussed above, this article is the result of research carried out by the Research Group HUM-311 of the Universidad de Málaga, which was organized into three subgroups of analysis according to the area being evaluated: language, mathematics and science. The subgroup responsible for the analysis of the area of science, which is the focus of this article, is composed of five re-

searchers specializing in the area of general teaching (four doctors and a researcher with a bachelor's degree in education) in addition to a doctor who is an expert in science teaching. The PISA tests in the area of science began to be investigated in 2008 through a preliminary in-depth study of the different theoretical frameworks related to both the PISA reports, and to relevant research in the area of educational competencies and the construct of Learning how to Learn in general. After analysing these theoretical frameworks, the Science group decided to move on to the analysis of the tests released by PISA in this area, followed by those appearing in the publications that are specified below: 2006 Spanish PISA Report (4 units), PISA 2006 Assessment Framework (16 units) and the items released from the 2000 and the 2003 PISA(13 units). For the analysis of the different

tests, a web format was used. Each researcher analyzed items from the PISA tests and they were then dumped into a common database for subsequent analysis in SPSS (figure 1 shows the common form and figure 2, the dating form). These analyses took the entire year of 2009, and there was continual discussion in order to achieve a triangulation of the results obtained in the analyses performed by each group member, subsequent to which the results were also checked with the researcher-expert in order to validate the findings. This took place either at in-person meetings held for this purpose (monthly) or by means of a virtual platform, based on the Moodle system, which favoured immediate and permanent contact among the six participating researchers (mainly through forums and by exchanging information in electronic format).

Analysis of Tests:						
<ol style="list-style-type: none"> 1. Title 2. Number of question: 3. Year: <ul style="list-style-type: none"> - Area: - Topic: - Process: - Competency required: 4. Correction criteria (sole answer or several possible responses) <ul style="list-style-type: none"> - Type of answer: No. of options: Maximum score (), partial points (), no points () 5. Skills (answer scored the highest points possible): 						
SKILLS						
	Reproduction	Application	Comprehension/ Reflection	Transfer	Heuristic	Communication/ Argumentative
LEVELS						
6. Observations						

Figure 1: Analysis basis form

PISA Sheet
Sole Group

0001: Commission

Please write answer here:

0002: Question Title

Please write answer here:

0003: Question Number

Please write answer here:

0004: Test year

Please write answer here:

0005: Type of answer

Choose the type of answer required for the question

Please write answer here:

- Open-ended
- Multiple choice
- Closed-ended

0006: Correction Criteria

Choose the type of correction criteria for the PISA question. If there is only one score (maximum or no points) or there are several possibilities (partial points, maximum points, etc.

Please choose one of the following:

- Sole Score (maximum or zero)
- Varied Score (maximum points, partial points)
- Comment on option here:

(Only answer this question if you answered Varied score (maximum points, partial points, etc.) on question no. 0006)

0007: Number of options

How many scoring options are there for the question, without counting no points. E.g. If partial points and maximum points are possible the number of options would be 2.

Please write the appropriate response for each entry:

(Only answer this question if you answered Varied score (maximum points, partial points, etc.) on question no. 0006)

0008: Skills (maximum score)

Please write the appropriate response for each entry:

	Level 0 (Absent)	Level 1	Level 2	Level 3
Reproduction				
Application				
Comprehension/Reflection				
Transfer				
Heuristic				
Communication/Argumentative				
Recover Information				

(Only answer this question if you answered Varied score (maximum points, partial points, etc.) on question no. 0006)

0009: Skills (2 point option)

Please write the appropriate response for each entry:

	Level 0 (Absent)	Level 1	Level 2	Level 3
Reproduction				
Application				
Comprehension/Reflection				
Transfer				
Heuristic				
Communication/Argumentative				
Recover Information				

Only answer this question if you answered Varied score (maximum points, partial points, etc.) on question no. 0006)

0010: Skills (3 point option)

Please write the appropriate response for each entry:

	Level 0 (Absent)	Level 1	Level 2	Level 3
Reproduction				
Application				
Comprehension/Reflection				
Transfer				
Heuristic				
Communication/Argumentative				
Recover Information				

0011: Observations

Figure 2: Dating Form

After triangulating the results, the analysis of the tests resulted in an organization which provided for both open discussion in the working group on each test in particular as well as the "reduction" of some characteristics to numerical gradients. Subsequently, a form was created with different fields in addition to a set of skills which we believed should be considered in the light of recent research (Monereo and Pozo, 2007). The following skills were identified in our analysis:

- **Reproduction:** involves mechanical and memory-based repetition of scientific knowledge.
- **Application:** this consists in applying the scientific knowledge previously learned to simple and/or known situations.
- **Comprehension/Reflection:** this involves the understanding of the phenomenon of science and reflection on the knowledge learned in this area.
- **Transference/Connection:** this means applying the knowledge learned to new situations, and connecting ideas, concepts, or scientific facts.
- **Heuristic:** solving the scientific problem requires the design of a 'plan', or the description of the steps that need to be followed to reach this solution.
- **Communication/argumentative:** the test requires well-argued reasoning to explain the scientific phenomenon concerned, and the communication of the scientific conclusion reached in written language.

These skills are required by PISA in the tasks to be carried out in each exercise. Therefore, since each exercise includes several questions or items often demanding more than one skill, it is logical that the number of skills identified exceeds the number of questions or items. Each skill established a priori was first valued based on how often it was found on the test and secondly, on its weight in the test. Additionally, a gradient of 1 to 3 was established to indicate whether the skill was dominant or secondary. After clearly defining these skills, in order to understand and analyze the tests in detail, the second step was to attempt to es-

tablish the parallels existing between these and the skills defined by PISA. The tests were then studied once again to examine the skills assessed by PISA in depth as well as their relationship with those considered by the research group. Under these considerations, the working group analysed the different tests again in order to clearly identify the previously defined skills. The analyses carried out and the findings obtained are summarized in the following section.

In summary, the methodological process followed in our study was carried out in the following phases:

- Study of the various theoretical frameworks on PISA.
- Development of an analysis form (shown above) on which each researcher recorded the data relating to the tests which he was to analyse: type of questions, skills required by the test, degree to which they were required, etc.
- After analysing the tests and extracting the related statistical data, as well as the specific information observed by each researcher in the tests (called "critical incidents" in this study), the different tests were again analysed in order to repeat and validate the results obtained. Throughout the whole process, meetings were frequently held in-person and, especially, virtually (using a virtual platform), in order to discuss the findings obtained.
- Lastly, the fruits of the research, the conclusions reached and the final results were included in a complete sector report on the area of science, of which this article is a synthesis.

Research Results

Skills found to be required on the tests

Based on the analysis of the tests carried out by the research group (table 1), it was found that the skill which was most often requested is that defined herein as 'reflection or comprehension' (96.08% of the total items), which means understanding the scientific phenomenon in question and reflecting on the scientific knowledge involved. In our analysis, it was found that this skill is often used on

data presented in the item itself, meaning that the student already knows the scientific phenomenon or process in question. Therefore in most of the tests a significant effort is not required from the students to solve the problem. In this regard, in the case of most of the items, the skill comprehension/ reflection does not extend beyond restating the information that is included in the text at the beginning of the item in order to stimulate students, without even requiring the student to relate the information from the text to what he already knows. In some cases the information is even underlined, as in the following test, which is presented for illustration purpose. For example:

Figure 3: Transit of Venus. Item 12.3 , of the 2006 theoretical framework

Question 12.3
In the sentence below several words have been underlined.
Astronauts predict that from the visual standpoint of Neptune, a Saturn transit will be able to be seen in front of the sun sometime in the present century.
Of the underlined words, which three are most useful for searching in a library or on internet to find out when this transit will occur?

Therefore, it can be concluded that although all of the tests involve comprehension, not all require deep reflection by students in order to solve the problem.

"Application" questions were also found to be very frequent (58.5% of the items analysed), this application generally being associated with contents found in the test. In this regard, it should be pointed out that, except on a few occasions, the tests analysed can be correctly answered by simply understanding the information provided in the same question

and then applying the knowledge which is usually present in the item itself.

On other occasions, this knowledge must be applied to a new problem, i.e. knowledge transfer is required, but this is only true in 32.36% of the items analysed. Likewise it was found that another of the higher skills defined by the research group, namely, the ability to design strategies to resolve the problem posed, or in other words, heuristic skills, are needed in a very low percentage of items in the tests analysed (12.78%).

In this regard, although transfer and heuristics/creation skills hardly appear in the tests analysed, it was decided that they would be included in the analysis since examples of questions requiring such skills can be found, as in the case of the test "Wind Power Plants" (Assessment Framework, 16.2, 2006, p. 183).

In this case, the student is provided with a number of conditions which are required to be met in relation to wind speed in order for wind turbines to actually generate energy. Students are prompted to choose the graph (from among four) which best represents the conditions described. We believe this question involves a transfer of knowledge to the problem posed and the design of a plan to develop the response (heuristic skills).

When faced with this question, students may decide to first make a list of all the variables that fall under the conditions (V1, V2, V3, W), then try to draw the graph, and then compare this graph with the answers provided. The question cannot ultimately be answered directly, but rather requires students to stop and think about how to answer it, and about what steps to take to find the answer, such as ordering and interpreting the data provided, etc.

Wind power plants

Many people believe that wind should replace petrol and coal as a source of energy for producing electricity. In the following image you can see some wind turbines, structures with blades that rotate with the wind. These rotations produce electricity using generators which are powered by the movement of the blades.



Question 16.2

The stronger the wind, the faster the wind turbine blades will rotate and the faster energy will be generated. However, in reality there is no direct relationship between the speed of the wind and the generation of electricity. Following are four operating conditions for the generation of electricity in a real wind power plant.

- The wind turbine blades begin to turn when the speed of the wind reaches the value V 1.
- The production of electricity reaches its maximum point (W) where the speed of the wind is V2
- For safety reasons, the plants are prohibited from rotating the blades at a higher speed than when the speed of the wind is V2.
- The blades stop spinning when the wind speed reaches the value V3.
-

Figure 4: "Wind power plants". from the Assessment Framework, 16.2, 2006, p.183

Furthermore, although, as indicated, the number of questions that can be answered without great difficulty using information already known by the student is not negligible, there are a few which simply call for "reproduction" of scientific knowledge and which as defined herein, is related to 'mechanical memorization'.

Specifically, such questions are only found in 16.67 % of the tests, which could be due to the way 'competence' is defined according to the PISA program itself, where reproduction is clearly ruled out in the assessment of competencies.

Table 1. Results of the PISA test analysis (% Maximum Score)

	RP	AP	RF	TR	HE	AR
Absent	67.73	34.31	1.96	49.02	72.55	54.90
1	12.75	28.31	12.75	12.75	1.96	4.90
2	3.92	12.75	26.47	17.65	2.94	9.80
3	0.00	17.65	56.86	1.96	7.84	17.65
NR	19.61	6.86	1.96	18.63	14.71	12.75
Present	16.67	58.5	96.08	32.36	12.78	32.35

Everything seems to indicate, therefore, that the tests are intended to measure the degree of acquisition of certain skills related to science, although not necessarily knowledge of science itself, but rather what we might call "logical thinking" Questions of this type are found relating to the Social Sciences and, of course, mathematics, in this case applied to biological or geological problems.

However, in the tests examined, an argument is not always demanded, and it seems that the only aspect of interest is the student's "applied" answer. It is the final result rather than the process which leads the student to choose one answer or another. In fact, even though argumentative skills are considered to be of most interest for the purpose of assessing complex thinking skills, only 32.35% of the items analysed require such skills.

Our study shows that argumentative skills are one of the key skills in the assessment of scientific competence, given that it is a skill which highlights students' complex thinking skills. In the tests analysed, when 'arguments are required, the question tends to involve a high difficulty level. However only one answer is considered to be valid, this answer being the one for which the most points are awarded. No consideration is made for the relations or the reflections that students may have made.

Open-ended questions, where argumentative skills are needed, require students to be used to reasoning and communicating their ideas, thereby moving away from mere 'recognition'.

Additionally, students have to know how to read comprehensively and to express themselves correctly in writing; all of which is highly important for an authentic assessment of competence.

What is PISA really trying to assess in its tests?

Although PISA proposes the measurement of three types of skills relating to what OCDE defines as "scientific competencies" (identifi-

cation of scientific questions, scientific explanation of phenomena and the use of scientific tests), it does not seem to pay special attention to the groups of higher or second order skills ('transfer, 'heuristic' and 'argumentative skills'), leading to a concept of competence which is tied almost exclusively to "knowing" as knowledge and neglecting the areas of 'know-how' and "know how to say" (DeSeCo) in the assessment of scientific competence. Therefore, although PISA is generally uninterested in scientific knowledge understood as the mere mechanic reproduction of information or recall of information from memory (which the group calls "reproduction"), we found that the second order skills indicated should be fostered more, as indicated by the PISA program itself in its theoretical foundations.

In this regard, it is important to highlight that PISA seems to veer toward multiple choice tests where argumentative skills, which in our opinion are essential in scientific competence, are not contemplated. Only certain activities include open-ended question requiring arguments. The rest of the problems only require choosing between multiple responses.

Additionally, when open-ended questions are used, the related correction criteria is closed, meaning that there is a limited set of possible answers which are considered to be correct. The way these answers are scored should be more focused on whether students were able to support their answer scientifically, a skill which is far more relevant for the assessment of competent learning. Also, we consider that student creativity and knowledge transfer should be assessed more favourably. As we have seen, there are very few occasions in which open-ended problems giving free rein to transfer skills are posed, and even less where there is an emphasis on heuristic skills.

Finally, there is a lack of activities that assess the skill of reading/interpretation of graphic images (e.g. sketches, drawings and diagrams of biology, geology, etc.), and a

greater use of written communication (e.g. writing a report, an assessment, etc.), scientific skills traditionally considered important in the teaching of science and now considered important in the development of scientific competence.

Relevance of the content of the tests

With respect to relevance, be reminded that reference is made to the concept of situated learning, which is considered to be more ideal for fostering learning. This means that activities must be connected to the interests or concerns of students in order to allow such learning to be developed. Learning competencies requires students to interpret and act on their immediate environment. Therefore, PISA is faced with a major problem with respect to the relevance of content: how is possible to connect with the interests and concerns of students using a standardized test which covers a large amount of subjects? The only way to make content relevant to the students is to know their concerns and interests, which is impossible to take into account on a test like PISA, in relation to which international standardization makes it difficult to create relevant contexts for socio-cultural realities

which are sometimes very different. It is important to emphasize that we are not claiming this is impossible in all cases, since for example, the use of tests based on events with global or a high media impact offer frameworks which are significant or relevant to all students, given the globalization of news.

However, although the content of the items is without doubt, updated and related to topics currently in discussion, PISA tends to raise such questions in an academic manner, wasting opportunities to raise ethical dilemmas or address current conflicts in the scientific field by merely focusing on the academic content of items (see "Cloning", PISA 2000-03).

This can also be seen in the type of answers envisaged by PISA in its questions. Of all the items analyzed, only 37% allow for the possibility of open-ended answers, closed and multiple choice answers accounting for 63% of

items (multiple choice answers being more prevalent in 41% of the items). The same occurs in relation to correction criteria: while in 81% of the questions, it is only possible to award either the total points assigned to a question or zero points, partial points can only be awarded in 19% of the items. It is clear that PISA does not contemplate the variety of perspectives and conflicts represented by science competence and pays more attention to the "verifiability" of academic knowledge.

PISA argues that if the student has knowledge of an issue in its immediate surroundings, it automatically becomes a topic which is relevant for him/her. However, this may be difficult to sustain, considering that we are surrounded by problems which are closely related to us, but not all of them are equally relevant or interesting.

Topics such as: "Physical Exercise", "The Greenhouse Effect", "Smoking", "Cloning", "Climate Change" and "Ozone", could be of some relevance given the widespread coverage of these topics in the media.

Others may be of interest to students given their relation to daily life. "Safe to drink", "Cavities", "Starlight", "Lip gloss", "Bread Dough", "Daylight" or "Chocolate". The other questions such as "Acid Rain", "Ultrasound", "Catalyst", "Bio-diversity", "Thorny Conduct" specifically in reference to the "Grand Canyon" (in Colorado, which is a bit too far away), are less likely to be relevant for boys and girls who are generally fifteen years old, although some may have a special interest in science or nature leading them to find the question more relevant. What is clear is that these can be no other than mere impressions for which there are no arguments regarding alleged relevance, since it is impossible to determine the degree of relevance of the content of the tests.

In summary, we find it difficult to pose questions which are relevant to the students through a standardized test for such a great variety of socio-cultural realities due to ignorance or the inability to connect with the con-

cerns and interests of such a diverse sample population. Consequently, it is impossible to have the knowledge necessary to provide for situated, meaningful, relevant and/or competent learning.

Additionally, it should be noted that the effort made by PISA to present current topics with a media impact in their items does not seem to be more than an update of their content and not a new way of addressing them. Therefore, on many occasions they end up seeking a degree of knowledge rather than competence in a in a field, as is traditionally done in our schools.

What is it that PISA does not pay attention to?

Although based on the evidence found in our study, PISA does appear to be assessing 'competencies' using the same approach as in DeSeCo (it being important to consider whether this can be done through standardised tests), following the analysis of the tests, is it possible to provide an overview of aspects to which the program appears to be paying less attention.

It seems obvious, based on the percentages of closed answers cited above, that argumentative and communication skills are not skills which PISA attempts to actively assess in their tests (this skill does not appear in 67.65% of the items analysed). Perhaps this is due to the difficulties that would arise in relation to the correction of the tests, which is linked to determining how to make sure corrections are objective, something impossible to guarantee if argument and communication are very prevalent.

Additionally, as we have seen, there are very few occasions in which open-ended questions giving free rein to the ability to create or to use heuristic skills are posed (this skill does not appear in 87.22% of the items analysed). It appears that PISA is more interested in monitoring knowledge of scientific theories and laws through closed-end questions) than in divergent thinking which is im-

plicit in the creative spirit and fundamental for the advancement of science.

Furthermore, one of the skills that PISA calls for is the 'search for and selection of scientific information'. It is a pity that this search is restricted only to the information provided in the question. In order for students to be able to demonstrate their competence to the utmost (remember that this has nothing to do with content) in relation to any aspect, it is imperative that all kinds of information are at their reach, including by means of a laptop connected to the internet network (which appears to have been proposed in PISA 2009, with a closed application).

Another noteworthy aspect of the PISA tests is the restriction of the problems to purely academic topics with little or no connection to cross-cutting themes. Consequently, although some activities pose questions related to health and the environment (never to other human rights), there are no questions in which students are expected to express their attitude; or to reflect on or consider ethical dilemmas, issues which are supposedly invariably linked to scientific activity.

In this regard, it is also noteworthy that certain items are included whose only purpose is to try to assess the students' attitudes towards science, although in a very superficial way. Curiously, the criteria for correcting these items have not been released, which to our understanding means that they do not count towards the final score on the test (see question 1.6 of "Safe to Drink" from March 2006).

Moreover, in relation to the content, it is possible that there is a cultural bias in the choice of the topics proposed, which are typical of a particular social class and far removed from local and minority cultures (see "Surgery with General Anesthesia" of March 2006, "Climate Change" of PISA 2000-03 and "The Catalyst" from the 2006 Framework).

The test designers' concern with avoiding any bias both in the content of the question and in the exactness of the responses causes

contents or issues to be avoided which are deemed to be problematic, in an attempt to flee from the "pollution" of opinions. Science, in its most positive sense, must respect certain processes and the findings obtained do not allow wrong answers. This exact knowledge and the assimilation of its methods can generate exacerbated rationalism and an attitude/belief toward scientific knowledge as being uncritical and lacking historical and social references, etc. to explain why science has drifted toward certain focuses of research.

No subjects of knowledge have been found in the tests which are open to debate, and which might lead students to doubt the options from which to choose or the answers required. However, this separation from problematic issues leads science and scientific knowledge to be seen as a metaphor of supreme truth, which for some is the new religion. Consequently, in this type of tests the examiners' strategy seems to be to eliminate "background noise" in the measurement, i.e. delete anything which is not strictly a scientific procedure. This is consistent with the way in which knowledge is compartmentalized in all educational curricula of the OECD.

The inclusion of the dimension "attitudes" in 2006 indicates that there has been a step forward in areas such as health and the environment, but nothing appears in relation to moral judgments regarding scientific activity.

Consequently, in the tests analysed, there appears to be a total lack of attention to the social and/or moral implications of the issues dealt with, for example, in aspects related to health, the environment or scientific research. In general, these absences are found in tests such as: "The Grand Canyon", "Smoking", "Physical Exercise", "Ozone", "Climatic Change", "The Greenhouse Effect" and "Cloning". In the above-mentioned tests, no reference is made to the impacts of tourism on the environment, the interests of multinationals in relation to certain diseases, ethics regarding reproduction, etc., where, as stated before, those paragraphs that are the cause of

moral, political, and ecological issues etc. are left out.

Conclusions

As analysed, PISA aims to establish an international common framework for the assessment of the performance of fifteen-year-old students, which is understood to be their level of competence. Under the concept of 'competencies', PISA has sought to promote the assessment of the 'usefulness' and 'application' of knowledge over reproduction in learning, by analysing and specifying the knowledge, skills and attitudes that are involved in learning scientific competence. However, in our analysis, we found a number of limitations with respect to the assessment of scientific competence in the terms defined by the PISA programme – as "complex know-how" -. In this regard, we consider it to be appropriate to encourage a more holistic and interdisciplinary approach to competencies, given the lack of correspondence between the concept of competencies posed by DeSeCo and the way in which PISA finally assesses these competencies by breaking them up into disciplines, capabilities, abilities and skills which are far removed from higher order skills such as transfer, heuristic and argument skills; effectively causing the concept of competence to lose its meaning. It is possible that the attempt to subject these tests to a conventional statistical analysis forces designers to use close-ended question exercises to make correction easier and avoid the subjectivity entailed by assessment.

Additionally, we consider that the full critical potential of the questions should be taken advantage of, bearing in mind that one striking aspect of the PISA tests seems to be the restriction of problems to topics that are purely academic, with little or no connection to cross-cutting and 'real life' topics. It is true that certain activities pose questions related to health and the environment, but without delving into them. Questions are not raised which thoroughly reveal the students' attitudes or pose problems leading students to reflect or

consider ethical dilemmas linked to scientific activity.

Additionally, we believe that allowing access to reference material for the purpose of answering the questions should be an essential resource for the assessment of competencies; from this standpoint, students should have all types of information within their reach.

In conclusion, in future research, one might consider why the assessment of competencies posed by PISA (through standardised testing) is or should be the most suitable for the purposes proposed by the assessment program itself. Some of the difficulties that have been observed throughout this article might prove insurmountable for a test of these features: the relevance of the contents of the tests, the objective correction of questions which might pose problems requiring arguments and/or the use of heuristic skills, etc. Therefore, we are in a position to conclude that there is a need for other models which are better suited to the assessment of educational competencies in holistic terms. Consequently, there are doubts as to whether the solution with respect to the assessment of competencies lies in perfecting the PISA test or in the model chosen for their assessment. Perhaps the countries involved should attempt qualitative approaches allowing for the observation of the behaviour of adolescents in natural situations in order to qualify global comprehensive statistical results given that the disappearance of these worldwide analyses does not seem to be on the horizon.

References

- Camaño, A. (2007). La evaluación PISA en Ciencias en 2006 en España e Iberoamérica. *Alambique. Did. Cien. Exper.* 47, 5-11.
- 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*, N° Extra. 295-331.
- Hargreaves, D.H. (2005). *Personalising learning 3: learning to learn & the new technologies*. London: Specialist Schools Trust.
- Harlen, W. (2002). Evaluar la alfabetización científica En el programa de la OCDE para la evaluación internacional de estudiantes (PISA). *Enseñanza de las Ciencias*, 20 (2), 209-216.
- Hatzinikita, V.; Dimopoulos, K. & Christidou, K. (2008). PISA test items and school textbooks related to science: A textual comparison. *Science Education*, 92 (4), 664 – 687.
- Hernández, F. (2006). El informe PISA: una oportunidad para replantear el sentido del aprender en la escuela secundaria. *Revista de Educación*, extraordinario PISA, 357-379.
- Ministerio de Educación (2006). Ley Orgánica 2/2006, de 3 de mayo, de Educación. *Boletín Oficial del Estado* n° 196 (4 de mayo), 17158-17207.
- Monereo, C. & Pozo, J.I. (2007). Competencias para (con) vivir con el siglo XXI. *Cuadernos de Pedagogía*, 370, 12-18.
- Organización para la Cooperación y el Desarrollo Económico (OCDE). (2002). *Conocimientos y aptitudes para la vida. Primeros resultados del PISA 2000 de la OCDE*. Madrid: Santillana-MEC.
- Organización para la Cooperación y el Desarrollo Económico (OCDE). (2002). *La definición y selección de competencias clave (DeSeCo). Resumen ejecutivo*. www.OECD.org/edu/statistics/deseco (www.deseco.admin.ch) (consultado el 10 diciembre 2009).
- Organización para la Cooperación y el Desarrollo Económico (OCDE). (2004). *Marcos teóricos de PISA 2003. Conocimientos y destrezas en Matemáticas, Lectura, Ciencias y Solución de problemas*, Madrid: Santillana-MEC.
- Organización para la Cooperación y el Desarrollo Económico (OCDE). (2006). *PISA 2006. Marco de la evaluación. Conocimien-*

tos y habilidades en Ciencias, Matemáticas y Lectura. Madrid: Santillana-MEC.

Oñorbe, A. (2008). Las pruebas de evaluación en Ciencias del proyecto PISA. *Alambique. Did. Cien. Exper.* 57, 41-52.


Psalidas, A., Apostolopoulos, C. & Hatzinikita, V. (2008). Investigating factors affecting students' performance to PISA Science items. *Journal of Engineering Science and Technology Review* 1, 90-97.


Puente Azcutia, J. (2008). PISA 2006: resultados españoles en Ciencias. *Alambique. Did. Cien. Exper.* 57, 12-22.


Notas


[1] Proyecto I+D (ref. SEJ-2007-66967), financiado por el Ministerio de Educación y dirigido por el profesor D. Ángel I. Pérez Gómez (Universidad de Málaga)


ABOUT THE AUTHORS / SOBRE LOS AUTORES


Gallardo-Gil, Monsalud (mongalgil@uma.es). PhD in Education Sciences from the Universidad de Málaga. She is the contact author for this article. She is in the research group "Educational Innovation and Assessment in Andalucía" (HUM-0311). Her main lines of research are: teacher training, ICT, educational assessment and teaching-learning processes in rural contexts. [Buscar otros artículos de esta autora en Google Académico / Find other articles by this author in Scholar Google](#) 

Fernández-Navas, Manuel (mfernandez1@uma.es). Researcher at the Universidad de Málaga. He has a Master's degree in Hearing and Language and Bachelor's degree in Educational Psychology. He is in the research group "Educational Innovation and Assessment in Andalucía" (HUM-0311). His lines of research are virtual teaching research-action, new technologies in education and educational assessment. He is currently doing his dissertation on virtual teaching at the university level. [Buscar otros artículos de este autor en Google Académico / Find other articles by this author in Scholar Google](#) 

Sepúlveda-Ruiz, María-Pilar (mdsepulveda@uma.es). Bachelor's degree in Pedagogy. Completed her PhD at the Universidad de Málaga, professor in the Department of Teaching and School Organization in the University. She is in the research group "Innovation and Educational Assessment in Andalucía" (HUM-0311). Her publications have focused on the initial training of teachers, her main concern throughout her professional career. [Buscar otros artículos de esta autora en Google Académico / Find other articles by this author in Scholar Google](#) 

Serván, María-José (servan@uma.es). PhD in Education Sciences, Assistant Professor in the Department of Teaching and School Organization at the Universidad de Málaga. She is in the research group "Innovation and Educational Assessment in Andalucía" (HUM-0311). Noteworthy are her publications on teacher training and educational assessment. [Buscar otros artículos de esta autora en Google Académico / Find other articles by this author in Scholar Google](#) 

Yus, Rafael (rafayus@telefonica.net). PhD in Sciences, tenured full professor of Natural Sciences and author of several dozens of books and hundreds of articles on science and pedagogy. He is in the research group "Educational Innovation and Assessment in Andalucía" (HUM-0311). He has been focusing his work on teaching science, cross-cutting themes, education in values and integrated curriculum. [Buscar otros artículos de este autor en Google Académico / Find other articles by this author in Scholar Google](#) 

Barquín, Javier (barquin@uma.es). e PhD in Educational Sciences, Associate Professor of the Department of Teaching and School Organization at the Universidad de Málaga. He is in the research group "Educational Innovation and Assessment in Andalucía" (HUM-0311). He has worked in the university since 1984, concentrating on matters relating to teachers and the use of ICT. [Buscar otros artículos de este autor en Google Académico / Find other articles by this author in Scholar Google](#) 

ARTICLE RECORD / FICHA DEL ARTÍCULO

Reference / Referencia	Gallardo-Gil, Monsalud; Fernández-Navas, Manuel; Sepúlveda-Ruiz, María-Pilar; Serván, María-José; Yus, Rafael & Barquín, Javier (2010). PISA and scientific competence: an analysis of the PISA tests in the Area of Science. <i>RELIEVE</i> , v. 16, n. 2, p. 1-17. http://www.uv.es/RELIEVE/v16n2/RELIEVEv16n2_6eng.htm
Title / Título	PISA and scientific competence: An analysis of the PISA tests in the Area of Science. [<i>PISA y la competencia científica: Un análisis de las pruebas de PISA en el Área de Ciencias</i>].
Authors / Autores	Gallardo-Gil, Monsalud; Fernández-Navas, Manuel; Sepúlveda-Ruiz, María-Pilar; Serván, María-José; Yus, Rafael & Barquín, Javier
Review / Revista	RELIEVE (Revista ELectrónica de Investigación y EValuación Educativa / <i>E-Journal of Educational Research, Assessment and Evaluation</i>), v. 16, n. 2.
ISSN	1134-4032
Publication date / Fecha de publicación	2010 (Reception Date : 2010 June 14; Approval Date : 2010 December 26; Publication Date : 2010 December 26).
Abstract / Resumen	<p><i>This work is part of an investigation focusing on the educational evaluation of the skill of Learning how to Learn, which addresses the issue of basic skills that students needs to develop to function throughout life in the contemporary social context. One purpose of the study is to analyze the potential of external and diagnostics evaluations to estimate the level of acquisition of these skills. Focusing on a methodology of qualitative and quantitative analysis of the PISA tests released in the specific area of Science (2000-2006), we conclude that these tests require scientific skills of low complexity, with marked presence of mere reproduction. In addition, we appreciate the gap between the concept of competence proposed DeSeCo and what PISA really evaluates according to the results obtained in the area of Science.</i></p> <p>El presente trabajo forma parte de una investigación centrada en la evaluación educativa de la competencia de Aprender cómo Aprender, en la que se aborda la temática de las competencias básicas que los estudiantes necesitan desarrollar para desenvolverse durante toda la vida en el contexto social contemporáneo. Uno de los propósitos del estudio es analizar el potencial de las evaluaciones externas y diagnósticas para estimar el grado de adquisición de estas competencias. Centrándonos en una metodología de análisis cualitativo y cuantitativo de las pruebas liberadas de PISA en el área específica de Ciencias (2000-2006), concluimos que estas pruebas demandan capacidades científicas de baja complejidad, con acentuada presencia de la mera reproducción. Además, valoramos la distancia existente entre el concepto de competencia propuesto en DeSeCo y lo que realmente parece evaluar PISA a la luz de los resultados obtenidos en el área de Ciencias.</p>
Keywords / Descriptores	<i>Assessment; Evaluation; Education; Skills; Competence; PISA; Sciences.</i> Evaluación; Educación; Competencias; PISA; Ciencias.
Institution / Institución	Departamento de Didáctica y Organización Escolar. Universidad de Málaga. (España).
Publication site / Dirección	http://www.uv.es/RELIEVE
Language / Idioma	Español & English version (Title, abstract and keywords in English)

RELIEVE

Revista ELectrónica de Investigación y EValuación Educativa
E-Journal of Educational Research, Assessment and Evaluation

[ISSN: 1134-4032]

© Copyright, RELIEVE. Reproduction and distribution of this articles it is authorized if the content is no modified and their origin is indicated (RELIEVE Journal, volume, number and electronic address of the document).

© Copyright, RELIEVE. Se autoriza la reproducción y distribución de este artículo siempre que no se modifique el contenido y se indique su origen (RELIEVE, volumen, número y dirección electrónica del documento).