Pre-service Primary School Teachers' Science Content Knowledge: an Instrument for its Assessment

José J. Verdugo[a], Joan J. Solaz-Portolés[b] and Vicente Sanjosé[b]

Corresponding Author: Joan.Solaz@uv.es
[a] Post-graduate Student at the University of Valencia, Spain.
[b] Department of Social and Experimental Sciences Education. University of Valencia, Spain

Keywords: scientific literacy, science content knowledge, pre-service primary school teachers, basic concepts, instrument validation.

International Journal of Innovation in Science and Mathematics Education, *24*(2), 37-51, 2016.

Abstract

An instrument to assess primary school science content knowledge in pre-service primary schoolteachers was elaborated and validated. A multiple-choice test with 50 items was initially proposed. Four content areas of science from the curriculum of Primary Education were considered: natural environment and its conservation, biological diversity, health education and the human body, matter and energy. A sample of 83 pre-service teachers from a Spanish university participated in the statistical validation process. External consistency (Test-Retest correlation), internal consistency (Kurder-Richardson's 20th formula), item difficulty and item discrimination indices were analysed as well as distractor selection. Some adjustments were made and a second shorter 30-item version was proposed. This new instrument was validated with four different samples amounted for a total of 756 Spanish pre-service primary teachers. Distributions of difficulty and discrimination indexes were analysed. Consistency among samples was very good. Remarkable results about participants' poor science content mastery and misconceptions were shown. Implications of these findings are discussed.

Introduction

Developed countries consider citizens' scientific education as essential for their development. However, it is not easy to determine what science education is, and what topics it should include. The science education benchmarks for the US (AAAS, 1993) introduced the general scientific topics considered as useful for all citizens in this country. Years later, in Lisbon 2000 the European Council agreed on the necessity of developing a knowledge-based economy. Renewing educational systems was carefully recommended in order to increase the amount of researchers in science, technology and mathematics (European Commission, 2004). In 2006, the Nuffield foundation promoted seminars on Science education in Europe. As a result, the *Science Education in Europe; Critical Reflections* report (Osborne & Dillon, 2008) argued against the 'Science for future Scientifics' goal and promoted the 'Science for all citizens' idea, in convergence with a previous Millar and Osborn's report (1998).

As a result of the debates generated, 'scientific literacy' has become one of the keywords in science education although its meaning is rather ambiguous. In this paper we considered the one provided by Hurd (1998): Scientific literacy "(…) *is seen as a civic competency required for rational thinking about science in relation to personal, social, political, economic problems, and issues that one is likely to meet throughout life*." (p. 410). In this vein, Gil and Vilches (2006) claimed that the purpose of science education should be to make citizens become scientific literate people, and to enable them to actively participate in decision making. For instance, in those socio-scientific issues on which there is not a clear experts' agreement but can seriously affect society.

Nowadays scientific literacy seems to be far from success in many countries, as revealed by different studies. In 2007, the European Commission report highlighted that traditional formal science education had a negative impact on children's attitudes towards science (Rocard, Csermely, Jorde, Lenzen, Walwerg-Henriksson, & Hemmo, 2007). International standardised studies, such as the Eurobarometer (Commission of the European Communities, 2005) and PISA (OECD, 2012), have also shown a comparative poor development of basic scientific competences in Secondary school students of some western countries, like Spain, despite the increasing budgets in Education.

Of course, there are many causal factors to explain these results. Improving teachers' education is probably an effective way to improve citizens' scientific literacy. Many teachers feel that they are not competent enough to be science literate and good science teachers (Morgan, 2012) although teachers' competence seems to vary across countries (see, for instance, studies done by Altun-Yalçin, Açişli & Turgut, 2011, or Cavas, Ozdem, Cavas, Cakiroglu, & Ertepinar, 2013).

The concept of scientific literacy (Laugksch, 2000) can be arranged in assessable components. Miller (1983) established three dimensions for scientific literacy and provided the concept with a rational structure: (a) an understanding of the rules and methods of science (i.e., the nature of science); (b) an understanding of key scientific terms and concepts (i.e., science content knowledge); and (c) an awareness and understanding of the impact of science and technology on society.

In this paper we focus on Miller's second dimension because it seems to be essential for the other two dimensions: nobody can teach what he/she does not properly know. Epistemic knowledge, involving the comprehension of the methods and rules of Science as a global enterprise, can hardly be elaborated if people have a lack of knowledge about what scientists are dealing with or what procedures they follow. Comprehending the real impact of science and technology on society seems very difficult if science content is not understood.

Teacher's content knowledge is a critical component in teaching as it affects what teachers teach and how they teach (Lewthwaite & MacIntyre, 2003). Literature has shown a strong relationship between science teachers' competence and their content knowledge (Alexander, 2000).

Teachers, especially Primary school teachers, need to have some specific scientific skills and basic science content knowledge (Baker, 1994) to promote children's curiosity and encourage them to explore their surrounding world and develop a useful knowledge for their daily lives (Smith & Neale, 1989). They need sufficient conceptual and procedural knowledge to develop suitable attitudes towards science teaching and learning, and to elaborate certain epistemological knowledge about the nature of science (Mellado, Blanco & Ruiz, 1998).

However, Primary school teachers usually lack enough science content knowledge (Appleton, 2003). In particular, in countries such as Australia or Spain, it has been reported that many elementary schoolteachers do not have an adequate competence for teaching science (Campanario, 1998; Cañal, 2000; Hackling, Peers, & Prain, 2007; MacDonald & Hoban, 2009). Lack of science content knowledge is often cited as the cause of teachers' inability to teach science effectively (Fleer, 2009) and it can lead their students to misconceptions (Kaptan & Korkmaz, 2001). Variations in teachers' science understanding have been identified as the main factor responsible for the differences in the quality of elementary science teaching (Shallcross, Spink, Stephenson, & Warwick, 2002).

Purpose of the Study
Surprisingly, there is a dearth of research addressed to study how to improve teacher's knowledge of science content (Fleer, 2009; Macdonald & Hoban, 2009). Reform actions addressed to improve teachers' science content knowledge should start with a reliable assessment to determine their conceptual obstacles, misconceptions and difficulties, as well as their correct knowledge. At the same time, reliable assessment requires appropriate and valid instruments.

Some instruments have been proposed and used with this purpose. The National Science Foundation validated a 9-item questionnaire to assess science factual knowledge in the US adult population (National Science Board, 2014). In Europe, similar studies have been reported (Commission of the European Communities, 2005; FECYT, 2015). Brossard and Shanahan (2006) elaborated a 31-item fill-in-the-blanks questionnaire based on scientific information frequent in mass-media. Rundgren, Rundgren, Tseng, Lin, and Chang (2010) designed the SLiM, a 50-item multiple choice instrument to assess scientific literacy in Secondary students. Science concepts were taken from textbooks and newspapers in Taiwan. Few studies have focused on primary school teachers' science content knowledge (Diamond, Maerten-Rivera, & Rohrer, 2013).

Up to our knowledge, there are not validated instruments to assess pre-service Primary School teachers' science content knowledge in Spain. Therefore, the main goal of this study is to design and to validate a tool for assessing Spanish primary school teachers' science content knowledge. The instrument should help identify pre-service Primary School teachers' misconceptions, as well as weaknesses and strengths in different science content knowledge areas. Such information could be used to develop better teacher training programmes.

Methodology

Instrument design and development
According to Jarrett, Ferry and Takacs (2012), multiple-choice instruments are suitable for data collection in educational surveys. In order to design a multiple-choice instrument for assessing Spanish primary school teachers' science content knowledge (excluding science skills), several steps were taken. First, a full revision of national curriculum documents was made to determine the main subject areas involved in elementary science in Spain. Four areas of curricular content (or "blocks") are distinguished in national curriculum documents: natural environment and its conservation ("Environment" onwards), biological diversity ("Life" onwards), health education and the human body ("Health"), and matter and energy ("Energy"). The content in these blocks is similar to the science content recommended by the AAAS' report (1989) and it has been taken into account in other instruments, such as in Laugksch and Spargo's (1996) one.

National curriculum documents for Spanish elementary science mentioned a long list of scientific concepts. In order to keep the minimum set of concepts covering the maximum knowledge, a set of well-known and widely used textbooks in Spain were analysed. We took into account the concept frequency and the relationship among the concepts in the main ideas of the lessons to determine the most outstanding concepts. A set of 50 scientific concepts were selected. These concepts were considered as "core" concepts in Primary school, and important for further scientific learning.

This set of basic concepts was taken to elaborate a first version of a questionnaire to assess primary science content knowledge (PSCQ-1 onwards). Items were elaborated trying to maintain their independence: answering a particular item correctly did not require knowledge of concepts involved in other items. Giving the right answer to the items did not involve procedural skills or epistemic knowledge but only definitions, knowing taxonomic or partonomic relations, application contexts, and relevant or specific functions of body organs. Items covered the aforementioned four curricular blocks.

Five answer options were provided for each item: a correct one, three distractors and an open one where participants had a blank space to write. Instructions to fill in the instrument informed that there was only a correct option in each item. They also told students to use the open option just and only if they considered the four options provided erroneous. In that way, we could obtain feedback about possible errors in the item design, reading obstacles or if the options provided were inadequate.

Validation Procedure for the initial questionnaire

The validation procedure followed two main steps. First, we performed a statistical validation. In order to assess the consistency and reliability of the questionnaire, a Test-Retest procedure was followed, with a two-week delay between the first and the second administration. Five intact groups (classroom groups formed by the university authorities) of pre-service Primary school teachers from a big public Spanish university participated in the study. We obtained valid data from 83 male and female students, aged 19-51, with an average of 22 years old. These subjects made up our sample. Participants were enrolled in “Science for Primary School-Teachers”, a compulsory subject. Each item was evaluated as correct (1 point) or incorrect (0 points).

The Test-Retest correlation was significant and strong (Pearson’s $r = 0.89$; $p< 0.001$). Taking into account that the instrument considered four different content blocks, the internal consistency was also high (KR-20 value, Test: 0.78, Re-test: 0.79) according to the George and Mallery’s criterion (2003, p. 231).

We explored the item Difficulty index (the proportion of correct answers). Averaging Test and Re-test, 27 items resulted “very easy” (average success over 0.85), or “easy” (success in the 0.66-0.85 interval), and 15 items reached intermediate success (0.46-0.65). In order to properly adjust the instrument difficulty to the target population, some authors like Garret (1958, cited in Rudramamba, 2004), and Summer (1959, also cited in Rudramamba, 2004) have suggested that a percentage about 25-30% of easy and difficult items should be included in any questionnaire together with other items (about 40-50%) of normal difficulty. According to this criterion, there were too easy items in PSCQ-1 and, in addition, more items with intermediate difficulty index should be included.

Item Discrimination indices were computed in a classical way (Kelly, 1939). We computed the difference $D_k = U_k – L_k$, where k is the particular item, U_k, is the proportion of correct answers to item k in the upper group (i.e. the 27% of students having the highest total scores), and L_k is the proportion of correct answers to item k in the lower group (i.e. the 27% of students with the lowest total scores). Items with D-values over 0.30 were considered good discriminants according to a usual criterion. On the other side, items having D-values less than 0.10 were considered non-discriminating items.

We also analysed the distractors proposed in each item. Non-functional distractors (NFD) have been defined by DiBattista and Kurzawa (2011) as wrong answer options in multiple-choice items obtaining a very low selection proportion, or a very low discrimination index. Thirty-four items (Sixty-eight per cent) had none or only one NFD in PSCQ-1. In addition, 6 items having two or three NDF revealed students’ misconceptions in science, and then were considered useful. Other 10 items have NFD to be revised or changed.

Participants spent around 45 min to fill in the questionnaire. After completion, ten participants were selected at random and interviewed to know their opinion on the easiness, or difficulty or suitability of the questionnaire, PSCQ-1. These students agreed on the benefit of reducing the fatigue with a shorter instrument, especially if the instrument was administered with other instruments (for instance, in concomitant assessment or in correlation studies).

Second, a group of 3 university teachers, experts in science education, analysed the 50 items proposed. They should judge whether they fit the Primary school curriculum or not. They suggested that some items should be re-written to facilitate comprehension. Two experts considered that 6 items from the "Environment" block included geography concepts instead of geology or environmental concepts. So they suggested removing them from the questionnaire.

This recommendation, together with the students' opinion, was taken into account at the end of the validation phase in order to propose a new shorter questionnaire (PSCQ-2). Several actions were performed:

a) Removing the "geography" items, according to the experts' suggestion.

b) Revising those items resulting too easy (percentage of correct answers over 0.85) or too difficult (percentage of correct answers less than 0.25) for participants.

c) Keeping the distribution of content blocks as near as possible to the distribution of these subjects in the Spanish national curriculum documents. This implied being aware of the relative percentage of items belonging to each content block: Environment: 22%, Life: 27%; Health: 31; Energy: 20%.

d) Revising (replacing or re-writing) the non-functional distractors.

From action (a), 6 items were removed and only 7 remained in the "Environment" content block. Therefore, and according to action (c), "Life" had to contain 8 items, "Health" 9 items and "Energy" only 6 items. Therefore, fifteen items of physics or chemistry content were removed from the "Energy" block, and one item was added to the "Health" block. According to action (b), items having too high or too low percentage of correct answers were removed first. Finally, non-functional distractors in the remaining items were revised and some were replaced.The resulting new version of the questionnaire, PSCQ-2, made up of 30 items, has been included in the Appendix 1. A new validation process for this shorten version was initiated.

Throughout two different years, one by one four samples (N_1= 189; N_2= 185; N_3= 185; N_4= 197; N_{tot}= 756) of pre-service teachers in three public Spanish universities participated in this new validation study. Participants were all enrolled in the same university Degree with a very similar curriculum in Science (and in Science Education) subjects. Each sample was made up of intact groups, randomly chosen from at least two different universities. The three samples were considered together in some analyses or independently in other statistical analyses (using SPSS 22.0). Results are shown and discussed below.

Results

Global results

Considering the four samples collapsed, the averaged percentage of correct answers reached 54.9% corresponding to a mean score of 16.5 out of 30, with a standard deviation (SD) of 4.2. Quartiles were defined by the scores 13.0; 16.5; 20.0. The content blocks obtained similar mean percentages of correct answers (from 49.4% in "Health", to 57.9% in "Life"). Table 1 shows the mean scores and standard deviations.

Table 1: Mean scores (and SD) obtained in PSCQ-2 by participants in the four samples.

	"Environment" (max=7)	**"Life"** (max=8)	**"Health"** (max=9)	**"Energy"** (max=6)	**Global** (max=30)
Sample1	3.9 (1.6)	4.6 (1.5)	4.5 (1.8)	3.3 (1.3)	16.3 (4.4)
Sample2	4.0 (1.6)	4.6 (1.4)	4.4 (1.8)	3.4 (1.3)	16.3 (4.2)
Sample3	4.0 (1.5)	4.7 (1.5)	4.5 (1.7)	3.4 (1.4)	16.6 (4.3)
Sample4	4.0 (1.5)	4.6 (1.5)	4.4 (1.6)	3.6 (1.3)	16.6 (4.1)
Total	4.0 (1.6)	4.6 (1.5)	4.4 (1.7)	3.4 (1.3)	16.5 (4.2)

A 3x4 Administration (sample1/sample2/sample3/sample4) X Block (Environment/ Life/ Health/ Energy) mixed ANOVA was conducted using normalized values (0-1) for block scores. The main effect of the 'Sample' factor was not significant (F< 1). Significance was obtained for the main effect of the Block factor, with a large effect size and high statistical power ($F(3,750) = 46.788$;

$p < .001$; $\eta^2 = 0.16$; P = 1.0). Thus, once the four samples were collapsed, there were intra-block differences. *Post-hoc* analysis showed significant differences in the "Health/ Environment" ($t(755) = -7.785$; $p < .001$), "Health/ Life" ($t(755) = -10.774$; $p < 0.001$), and "Health/Energy" ($t(755)= -8.601$; $p < 0.001$) comparisons. Finally, and more interesting, there was not significant "Sample x Block" interaction (F<1). Therefore inter-block differences were statistically similar in the four samples.

Regarding between-block correlations (Table 2), all the Pearson's coefficients reached significance ($p < 0.01$), although the values were not high. It should be kept in mind that differences in the science content among blocks were maximised in the shorten version, PSQC-2.

Table 2: Between-blocks correlation values (p < 0.01 in all cases)

Pearson's coefficient	Energy	Health	Life
Health	0.27		
Life	0.28	0.36	
Environment	0.28	0.29	0.24

External consistency

The four different samples obtained very similar results in PSQC-2, in coherence with the ANOVA conducted before. Averages and standard deviations were very close, as shown in Table 1. There were not significant differences between any couple of samples, according to the t-tests conducted ($p > 0.32$ in any case)

Internal consistency

The global consistency of PSCQ-2 was KR-20 = 0.65, meaning a 'questionable consistency' according to the George and Mallery (2003; p. 231) 'thumb' criterion. Additional analyses showed that the KR-20 value decreased when any single item was removed. Compared to the first and longer version PSQC-1, internal consistency diminished in this second version probably due to the well-known effect of the number of items on this indicator (especially when non-discriminating items were removed, as in this case).

Therefore, PSCQ-2 is far from a "single construct" instrument. Each of the 30 items was elaborated to be independent in order to cover as much diverse science topics as possible. The diverse knowledge implied in different items seems to be weakly related in subjects' memory, and then item-to-item correlations were not strong.

Of course, none of the content blocks considered individually obtained acceptable consistency, probably due to the low quantity of items involved (KR-20 < 0.50 in all cases). Therefore, the PSCQ-2 cannot provide a reliable assessment for each of the single elementary science areas considered.

Item difficulty indices

After the revision made to the first version of the questionnaire PSCQ-1, we explored the difficulty of the 30-items included in the shorter version PSCQ-2. We expected that most items would have a "normal" (intermediate) difficulty. Table 3 shows the distribution of the item difficulty index (or percentage of success) in PSCQ-2.

Table 3: Distribution of the item difficulty index in PSCQ-2.

Interval	Amount of items
0-0.25 (very difficult)	2
0.26-0.45 (difficult)	5
0.46-0.65 (intermediate)	14
0.66-0.85 (easy)	7
0.86-1 (very easy)	2

Thus, 47 percent of items reached an intermediate level of success, in agreement to Garret's (1958) and Summer's (1959)' thumb' criteria (both cited in Rudramamba, 2004). The Appendix 1 shows the percentage of global success in each of the items in PSCQ-2 (see also Table 4 below). In comparison to the first version, PSCQ-1, the mean success decreased a little bit. Changes made produced a kind of "transfer" from "easy" items to "normal" items, so 50% of the items in PSCQ-1 had percentages of success in the 0.66-0.85 interval, but in the second version, PSCQ-2, near 50% of the items have percentages of success in the 0.46-0.65 interval.

Item discrimination indices

In addition, we calculated the discrimination indices, D, in a classical way. Thus, we considered the upper 27 per cent subgroup (the subjects with the 27 per cent highest global scores) and the lower 27% subgroup in the sample. Then, for each item, we computed the difference $D_k = U_k – L_k$, where k is the particular item, U_k, is the proportion of correct answers to item k in the upper group, and L_k is the proportion of correct answers to item k in the lower group (Kelly, 1939). Answer options with positive D-values are those selected by the high-score group of students more frequently than by low-score group. Items with D-values over 0.30 are considered good discriminants and D-values between 0.10 and 0.30 are considered fair discriminants. On the other side, D-values near to 0 indicate low or no discrimination between high and low skilled students according to that instrument. In PSCQ-2, there was not an item with low discrimination value ($D < 0.10$) whereas 20 out of 30 appeared to be good discriminants ($D > 0.30$). The remaining items had discrimination indexes between 0.10 and 0.30. All the item-total (biserial-point) correlations reached significance ($p< 0.01$).

Discrimination indices are related to difficulty indices as very low or very high percentages of success are not compatible with high discrimination indexes. Table 4 shows the distribution of items in PSCQ-2 according to their discrimination power and difficulty indexes (see Table 3).

Table 4: Classification of items in PSCQ-2 according to the difficulty and the discrimination indices.

Difficulty Discrimination	Very easy	Easy	Intermediate	Difficult	Very difficult
High (D>0.30)	---	1, 2, 7, 10, 16, 28	3, 9, 11, 13, 15, 17, 18, 20, 22, 26, 27	5, 24, 30	---
Fair (0.10<D<0.30)	8, 19	25	4, 14, 29	6, 12,	21, 23
Low (D< 0.10)	---	---	---	---	---

Most of the items considered easy or difficult had intermediate discrimination indices showing that they were easy or difficult for most students, no matter their global performance. Some "normal" items (i.e. items with an intermediate level of difficulty) were good discriminants. That means that they did not seem "normal" to all students, but easy for the high performance sub-group, and difficult for the low performance sub-group.

Pre-service Primary teachers' content knowledge

The analysis of distractors may be of great educational interest. Some items obtained low levels of success (less than 50%) so revealing a poor mastery, but not a clear conceptual error. In these items, at least two distractors obtained a high frequency compared to the one in the correct option. Items # 5 and #14 defined this group. However, in item #5 two distractors reached similar success to the correct option, while in item #14 the correct option was nearly of 50% of success and the other three distractors reached similar percentages.

Other items suggested troubling misconceptions in future teachers or, at least wrong learning. This was found because in these items typical and well-known conceptual errors were associated to distractors specifically designed to this purpose (and also included in the first version, PSCQ-1). Obviously, we expected such misconceptions would be present in a small percentage of future primary school teachers, but this was not the case in some topics. In these items only one distractor obtained a high frequency, compared to the one in the correct option. This was the case of items #4 #6, #12, #13, #21, #23, #24, #27 and #30, for which the wrong answers concentrated on a single distractor which reached a relative high frequency. In items #6, #12, #21, #23, #24 and #30 the most chosen distractor overpassed the percentage of correct answers (see Table 5) thus suggesting several learning problems.

Table 5: Participants' most frequent errors. Percentages of selection are given in brackets.

Item	Question	Frequent Error(s)	Correct Answer
#4	Which of the following statements is correct?	Minerals, like marble and granite, are very useful for construction (32%)	Rocks are made of minerals (49%)
#6	What is the name given to the groups formed by gas, interstellar dust and thousands or millions of stars?	Nebulas (50%)	Galaxies (34%)
#12	What is the name given to the leaves that form the calyx of a flower?	Petals (48%)	Sepals (29%)
#13	What is the function of the phloem in a plant	Carrying raw sap from the roots to the rest of the plant (31%)	Carrying elaborated sap from the leaves to the rest of the plant (46%)
#21	In the excretory system, what duct connects to the outside of the body?	Anus (68%)	Urethra (18%)
#23	Which organ produces bile?	Pancreas (57%)	Liver (23%)
#24	What nutritional substances, present in cereals and legumes, provide us with energy?	Proteins (52%)	Carbohydrates (30%)
#27	Which of the following statements about renewable energy sources is correct?	Thermal solar energy (…) generates electricity by means of photovoltaic solar panels (37%)	They are considered inexhaustible energy sources (48%)
#30	What colour would be obtained if all the colours of the rainbow were mixed?	Black (58%)	White (36%)

Conclusions and implications

The main objective of this work was designing and validating an instrument to assess elementary science content knowledge in pre-service Primary School teachers. The instrument was elaborated after analysing some national curriculum documents and bestselling textbooks. This analysis resulted first in a set of 50 most-relevant selected concepts in elementary science. Each single item was elaborated using just one different concept. At the end of the validation study we obtained suitable Test-Retest consistency and enough reliability for the intrument (PSCQ-1). However, long completion times and some other indicators provided evidence of the need to shorten the questionnaire. Some changes and adjustments were made so that it fit primary teachers' population better. Changes introduced resulted in a second, shorter version, PSCQ-2, made up of 30 multiple-choice items. Four different samples participated in a validation study of this second version of the questionnaire. Results were very stable and suggest that PSCQ-2 would be a useful instrument to assess Spanish pre-service teachers' elementary science content knowledge. Of course minor additional adjustments could be performed so that it fits this population better.

On the other hand, we expected primary school teachers would have better science content knowledge, especially when only primary concepts involved in the Spanish national curriculum were considered in the instrument. However, we should take into account that the present instrument has been adapted to the Spanish pre-service teacher population in some way, thus creating an "average shift" to central scores (with average success rate of 55%). This poor result is in line with past similar studies on teachers' science knowledge developed in Spain (Campanario, 1998). Moreover, the lack of basic scientific knowledge is one the (self-)perceived obstacles by Spanish pre-service primary teachers (Cortés, Gándara, Calvo, Martínez, Ibarra, Arlegui, & Gil, 2012). However, Spanish results differ from the ones obtained in other countries such as Turkey (Cavas, Ozdem, Cavas, Cakiroglu & Ertepinar, 2013) or Taiwan (Chin, 2005). These authors obtained satisfactory results (average success rate of 64% in Turkey, and 75% in Taiwan) in similar studies about teachers' scientific literacy including science content knowledge.

Some items have highlighted possible pre-service teachers' misconceptions in elementary science content. For instance, as shown in Table 5, almost 50 percent of the subjects considered "petals" as the name of the leaves forming the calyx in a flower (the correct option, "sepals" obtained 29 percent of choice) or they said that blending the rainbow colours together results in colour "black" (58%) instead of "white" (36 percent chose the correct option). Near 68 percent wrongly chose the "anus" as a part of the excretory system connecting to the outside of the body (only 18 percent of the answers were correct: "urethra").

Science misconceptions (and wrong learning) in teachers have been revealed in many studies conducted in different countries. Our results are consistent with previous evidence of pre-service teachers' misconceptions and conceptual errors in different science subjects obtained in previous studies (Bisard, Aron, Francek, & Nelson, 1994; Keles, Ertas, Uzun, & Cansiz, 2010; Kikas, 2004; Schoon & Boone, 1998; Trundle, Atwood, & Christopher, 2002). In a recent review, Kind (2014) analysed studies about teacher's subject-based misconceptions in physics, chemistry, space and Earth sciences and biology. She concluded that teachers show the same misconceptions than their students. However, she also obtained evidence of improvement of teachers' science content knowledge as they mature and receive further in-service education. What is the reason why? Probably, most pre-service teachers enrol in the university degree with the same misconceptions as the rest of the population. At university, science subjects do not help them revisit their school science knowledge. Instead, it develops further science subjects. Therefore, poor or wrong prior learning is not re-examined and it remains for a long time.

It is always difficult to determine how to overcome such misconceptions but probably active learning methods (Michael, 2006) such as giving the students more opportunities to deal with experimental work, and engaging in inquiry learning could solve part of the problems. Of course, contrasting these conjectures requires procedural and pedagogical content knowledge in science to be assessed as well in pre-service teachers.

These facts should warn us of the way Primary School-teachers are being trained in elementary science at Spanish universities.

Acknowledgments

We would like to thank the referee for his/her kind and encouraging suggestions addressed to improve the manuscript content, but also to make us easier overcoming our foreign language limitations.

References

Aiken, L. R. (2003). *Psychological testing and assessment (11th ed.)*. Boston: Allyn and Bacon.

Alexander, P. (2000). *An investigation into teacher and student teacher confidence in their own understanding and abilities to teach science and technology in primary schools in Northern Ireland* (Doctoral dissertation). Queen's University of Belfast, Belfast, Northern Ireland.

Altun-Yalçin, S., Açişli, S., & Turgut, Ü. (2011). Determining the levels of pre-service science teachers' scientific literacy and investigating effectuality of the education faculties about developing scientific literacy. *Procedia Social and Behavioral Sciences*, *15*, 783–787. doi:10.1016/j.sbspro.2011.03.185. Retrieved October 7, 2015, from http://www.sciencedirect.com/science/journal/18770428

American Association for the Advancement of Science [AAAS] (1989). *Science for all Americans*. Retrieved October 7, 2015, from http://www.project2061.org/publications/sfaa/online/sfaatoc.htm

American Association for the Advancement of Science [AAAS] (1993). *Benchmarks for science literacy*. Retrieved October, 7, 2015, from http://www.project2061.org/publications/bsl/online/index.php

Appleton, K. (2003). How do beginning Primary School teachers cope with Science? Toward an understanding of Science teaching practice. *Research in Science Education*, *33*(1), 1-25. doi: 10.1023/A:1023666618800

Baker, R. (1994). Teaching science in primary schools: What knowledge do teachers need? *Research in Science Education*, *24*(1), 31-40. doi: 10.1007/BF02356326

Bisard, W. J., Aron, R. H., Francek, M. A., & Nelson, B. D. (1994). Assessing selected physical science and earth science misconceptions of middle school through university preservice teachers. *Journal of College Science Teaching*, *24*(1), 38–42.

Brossard, D. & Shanahan, J. (2006). Do they know what they read? Building a scientific literacy measurement instrument based on Science media coverage. *Science Communication*, 28, 47-63.

Campanario, J. M. (1998) ¿Quiénes son, qué piensan y qué saben los futuros maestros y profesores de ciencias?: Una revisión de estudios recientes. *Revista Interuniversitaria de Formación del Profesorado*, *33*, 121-140.

Cañal, P. (2000). El conocimiento profesional sobre las ciencias y la alfabetización científica en primaria. *Alambique*, *24*, 46-56.

Cavas, P. H., Ozdem, Y., Cavas, B., Cakiroglu, J., & Ertepinar, H. (2013). Turkish pre-service elementary science teachers' scientific literacy level and attitudes toward science. *Science Education International*, *24*(4), 383-401. Retrieved February 17, 2016, from: http://files.eric.ed.gov/fulltext/EJ1022326.pdf

Chin, C. C. (2005). First-year Pre-service Teachers in Taiwan—Do they enter the teacher program with satisfactory scientific literacy and attitudes toward science? *International Journal of Science Education*, *27*(13), 1549-1570.

Commission of the European Communities (2005). *Europeans, Science and Technology. Special Eurobarometer 224*. Brussels. Retrieved January 9, 2016 from: http://ec.europa.eu/COMMFrontOffice/PublicOpinion/index.cfm/Survey/getSurveyDetail/yearFrom/2004/yearTo/2007/surveyKy/447/p/3.

Cortés, A.L., Gándara, M., Calvo, J.M., Martínez, M.B., Ibarra J., Arlegui, J. & Gil, M. J. (2012). Expectativas, necesidades y oportunidades de los maestros en formación ante la enseñanza de las ciencias en la Educación Primaria. *Enseñanza de las Ciencias*, *30*(3), 155-176.

Croker, L. & Algina, J. (2006). *Introduction to classical and modern test theory*. Ohio: Cengage Learning.

Diamond, B. S., Maerten-Rivera, J., Rohrer, R. & Lee, O. (2013). Elementary teachers' science content knowledge: Relationships among multiple measures. *Florida Journal of Educational Research*, *51*, 1-20. Retrieved March 14, 2016, from http://feraonline.org/fjer/2013/Diamond.51%281%29.pdf

DiBattista, D. & Kurzawa, L. (2011). Examination of the Quality of Multiple-choice Items on Classroom Tests. *The Canadian Journal for the Scholarship of Teaching and Learning*, *2*(2), Article 4. Retrieved October 16, 2015, from http://dx.doi.org/10.5206/cjsotl-rcacea.2011.2.4.

European Commission (2004). *Europe needs More Scientists: Report by the High Level Group on Increasing Human Resources for Science and Technology*. Brussels. Retrieved May 9, 2016, from: https://ec.europa.eu/research/conferences/2004/sciprof/pdf/conference_review_en.pdf

Fleer, M. (2009). Supporting scientific conceptual consciousness or learning in 'a roundabout way' in play-based contexts. *International Journal of Science Education*, *31*(8), 1069-1089. doi: 10.1080/09500690801953161

FECYT: Fundación Española Ciencia y Tecnología 2014 (2015). *Percepción Social de la Ciencia y la Tecnología en España*. Madrid. Retrieved January, 12, 2016 from: http://www.fecyt.es/es/publicacion/percepcion-social-de-la-ciencia-y-la-tecnologia-2014.

George, D. & Mallery, P. (2003). *SPSS for Windows step by step: A simple guide and reference*. Boston: Allyn & Bacon.

Gil, D. & Vilches, A. (2006). Educación ciudadana y alfabetización científica: mitos y realidades. *Revista Iberoamericana de Educación*, *42*, 31-53. Retrieved November 23, 2015, from http://www.rieoei.org/rie42a02.pdf

Hackling, M., Peers, S. & Prain, V. (2007). Primary Connections: Reforming science teaching in Australian primary schools. *Teaching Science*, *53*(3), 12-16. Retrieved February 10, 2016, from http://ro.ecu.edu.au/cgi/viewcontent.cgi?article=2445&context=ecuworks

Hurd, P. D. (1998). Scientific Literacy: New minds for a changing world. *Science Education*, *82*(3), 407-416. doi:10.1002/(SICI)1098-237X(199806)82:3<407::AID-SCE6>3.0.CO;2-G

Jarrett, L., Ferry, B. & Takacs, G. (2012). Development and validation of a concept inventory for introductory-level climate change science. *International Journal of Innovation in Science and Mathematics Education*, *20*(2), 25-41. Retrieved February, 14, 2016 from: http://openjournals.library.usyd.edu.au/index.php/CAL/issue/view/490

Keles, Ö., Ertas, H., Uzun, N. & Cansiz, M. (2010). The understanding levels of preservice teachers' of basic science concepts' measurement units and devices, their misconceptions and its causes. *Procedia Social and Behavioral Sciences*, *9*, 390-394.doi:10.1016/j.sbspro.2010.12.170. Retrieved February 14, 2016, from http://www.sciencedirect.com/science/journal/18770428

Kind, V. (2014). Science teachers' content knowledge. In Hamsa Venkat, Marissa Rollnick, John Loughran and Mike Askew (Eds), *Exploring Mathematics and Science Teachers' Knowledge. Windows into teacher thinking* (pp. 15-28). New York: Routledge, Taylor and Francis Group.

Kaptan, F. & Korkmaz, H. (2001). Primary school preservice teachers' misconceptions about heat and temperature in science teaching. *Hacettepe University Journal of Education*, *21*, 59-65. Retrieved February 14, 2016, from http://www.efdergi.hacettepe.edu.tr/yonetim/icerik/makaleler/1030-published.pdf

Kelly, T. L. (1939). Selection of upper and lower groups for the validation of test items. *Journal of Educational Psychology*, *30*(1), 17-24. doi: http://dx.doi.org/10.1037/h0057123

Kikas, E. (2004). Teachers' conceptions and misconceptions concerning three natural phenomena. *Journal of Research in Science Teaching*, *41*(5), 432-448. doi: 10.1002/tea.20012

Laugksch, R. (2000). Scientific Literacy: A conceptual overview. *Science Education*, *84*(1), 71-94. doi: 10.1002/(SICI)1098-237X(200001)84:1<71::AID-SCE6>3.0.CO;2-C

Laugksch, R.C. & Spargo, P. (1996). Construction of a paper-and- pencil test of basic scientific literacy goals recommended by the American Association for the Advancement of Science. *Public Understanding of Science*, *5*(4), 331–359. doi:10.1088/0963-6625/5/4/003

Lewthwaite, B. & MacIntyre, B. (2003) Professional knowledge, interest and self-efficacy: a vignette study. *STER papers: science and technology education research papers*, 161-188. Retrieved February 17, 2016, from http://researchonline.jcu.edu.au/34485/3/34485%20Lewthwaite%20and%20MacIntyre%202003%20Accepted%20version.pdf

Lord, F. M. (1952). The relation of the reliability of multiple-choice tests to the distribution of item difficulties. *Psychometrika*, *17*(2), 181-194. doi: 10.1007/BF02288781

Macdonald, D. & Hoban, G. (2009) Developing Science Content Knowledge Through the Creation of Slowmations. *The International Journal of Learning*, *16*(6), 319-330. Retrieved January 21, 2016, from http://ro.uow.edu.au/cgi/viewcontent.cgi?article=1108&context=edupapers

Mellado, V., Blanco, L.J., & Ruiz, C. (1998). A framework for learning to teach Science in initial primary teacher education. *Journal of Science Teacher Education, 9* (3), 195-219.

Michael, J. (2006). Where's the evidence that active learning works? *Advances in Physiology Education*, *30*(4), 159-167. Retrieved November, 13, 2015 from: http://advan.physiology.org/content/30/4/159.full-text.pdf+html

Millar, R. & Osborne, J. (1998). *Beyond 2000: Science education for the future: A report with ten recommendations*. London, UK: King's College London, School of Education.

Miller, J. D. (1983). Scientific literacy: A conceptual and empirical review. *Daedalus*, *112*, 29-48.

Morgan, A. (2012). 'Me as a Science Teacher': Responding to a Small Network Survey to Assist Teachers with Subject-Specific Literacy Demands in the Middle Years of Schooling. *Australian Journal of Teacher Education*, *37* (6). doi: http://dx.doi.org/10.14221/ajte.2012v37n6.6

National Science Board [NSB] (2014). *Science and Technology: Public attitudes and understanding*. In *Science and Engineering Indicators 2014* (7.20-7.23). Arlington VA: National Science Foundation (NSB 14-01). Retrieved October 15, 2015, from http://www.nsf.gov/statistics/seind14/index.cfm/chapter-7/c7h.htm

OECD (2012). *PISA 2012 Results*. Retrieved April, 7, 2016 from: http://www.oecd.org/pisa/keyfindings/pisa-2012-results.htm

Osborne, J. & Dillon, J. (2008). *Science education in Europe: Critical reflections* (Vol. 13). London: The Nuffield Foundation. Retrieved May 5, 2016, from http://efepereth.wdfiles.com/local--files/science-education/Sci_Ed_in_Europe_Report_Final.pdf

Rocard, M., Csermely, P., Jorde, D., Lenzen, D., Walwerg-Henriksson, H. & Hemmo, V. (2007). *Science Education Now: A Renewed Pedagogy for the Future of Europe*. European commission. Community Research. Retrieved November 15, 2015, from http://ec.europa.eu/research/science-society/document_library/pdf_06/report-rocard-on-science-education_en.pdf

Rudramamba, B. (2004). *Problems of teaching*. New Delhi: A.P.H. Publishing Corporation.

Rundgren, C.J., Rundgren, S.N. C., Tseng, Y.-H., Lin, P.L. & Chang, C.Y. (2010). Are you SLiM? Developing an instrument for civic scientific literacy measurement (SliM) based on media coverage. *Public Understanding of Science*, *20*(10), 1–15. doi: 10.1177/0963662510377562. Retrieved October 29, 2015, from https://www.researchgate.net/profile/Chun_Yen_Chang/publication/247156553_Are_you_SLiM_Developing_an_instrument_for_civic_scientific_literacy_measurement_%28SLiM%29_based_on_media_coverage/links/545972bd0cf2bccc4912bcee.pdf

Schoon, K. J. & Boone, W. J. (1998). Self-efficacy and alternative conceptions of science of preservice elementary teachers. *Science Education*, *82*(5), 553-568. doi: 10.1002/(SICI)1098-237X(199809)82:5<553::AID-SCE2>3.0.CO;2-8

Shallcross, T., Spink, E., Stephenson, P. & Warwick, P. (2002). How primary trainee teachers perceive the development of their own scientific knowledge: Links between confidence, content and competence? *International Journal of Science Education*, *24*(12), 1293-1312. doi: 10.1080/09500690110110106

Smith, D. C. & Neale, D. C. (1989). The construction of subject matter knowledge in primary science teaching. *Teaching and Teacher Education*, *5*(1), 1-20. doi: 10.1016/0742-051X(89)90015-2

Trundle, K. C., Atwood, R. K. & Christopher, J. E. (2002). Preservice teachers' conceptions of moon phases before and after instruction. *Journal of Research in Science Teaching*, *39*(7), 633-658. doi: 10.1002/tea.10039

Appendix 1. The instrument.

The percentage of subjects choosing the correct option (highlighted with *) is shown. Blank answers are not considered.

QUESTION	ANSWER OPTIONS (*correct answers)
1. What is the movement of the Earth around the Sun called? (70%)	a) Rotation b) Precession c) Revolution* d) Circumference
2. How long does it take the Moon to orbit the Earth? (68%)	a) 21 days b) 28 days* c) 365 days d) 7 days
3. Climate is… (63%)	a) The meteorological features occurring in a specific region over a long period of time* b) The set of atmospheric phenomena occurring in a specific time and place c) The condition of the atmosphere for a short period of time (less than 30 days) d) The meteorological conditions produced by changes in atmospheric pressure
4. Which of the following statements is correct? (49%)	a) Minerals, like marble and granite, are very useful for construction b) Rocks have a low value for human beings due to their few applications c) Minerals are made of rocks d) Rocks are made of minerals*
5. What kind of device is used for measuring wind speed? (36%)	a) Weathercock b) Pluviometer c) Anemometer* d) Barometer
6. What is the name given to the groups formed by gas, interstellar dust and thousands or millions of stars? (34%)	a) Nebulas b) Constellations c) Galaxies* d) Planetary systems
7. What part of a volcano is the conduit? (76%)	a) Underground pool where magma resides temporary b) Orifice connecting volcano to the exterior c) Pipe by which magma flows up* d) Rocks and solid material formed when lava gets cold
8. Which of the next animal groups are invertebrates? (86%)	a) Birds b) Fishes c) Insects* d) Amphibians
9. Living things can be arranged in levels of complexity. When some tissues co-work to perform the same function, what do they form? (63%)	a) Organs* b) Apparatus c) Molecules d) Interwoven systems
10. What vital function do flowers contribute to in plants? (73%)	a) Respiration b) Reproduction* c) Nutrition

	d) Interaction
11. In addition to water and sun, what gas do plants need to carry out the photosynthesis? (62%)	a) O_2 (oxygen) b) CO (carbon monoxide) c) CO_2 (carbon dioxide)* d) N_2 (nitrogen)
12. What is the name given to the leaves that form the calyx of a flower? (29%)	a) Sepals* b) Stamens c) Petals d) Corolla
13. What is the function of phloem in a plant? (46%)	a) Absorbing rays from solar light b) Carrying elaborated sap from the leaves to the rest of the plant* c) Carrying raw sap from the roots to the rest of the plant d) Absorbing water and mineral salts from the soil
14. What biological kingdom do viruses belong to? (49%)	a) Fungi b) Monera c) Protista d) No kingdom*
15. Which of the following statements about differences between plant cells and animal cells is correct? (55%)	a) Animal cells have mitochondria and plant cells do not b) Plant cells contain organelles called chloroplasts that animal cells do not* c) Plant cells do not have cellular membrane and animal cells do d) There are no differences. All living organism are made up of the same cells
16. Where is the exchange between O_2 and CO_2 produced in the respiratory system? (67%)	a) Bronchi b) Bronchioles c) Pulmonary alveolus* d) Trachea
17. Which organ is responsible for filtering the blood to remove waste substances? (64%)	a) Stomach b) Liver c) Kidneys* d) Urinary bladder
18. Which blood component has the function of carrying oxygen? (56%)	a) Plasma b) Platelets c) White cells d) Red cells*
19. What is the name given to the cells resulting from fertilization? (86%)	a) Gamete b) Ovule c) Spermatozoon d) Zygote*
20. Most nutrients pass into the blood in… (51%)	a) The large intestine b) The small intestine* c) The oesophagus d) The stomach
21. In the excretory system, what duct connects to the outside of the body? (18%)	a) Ureter b) Anus c) Urethra* d) Urinary bladder
22. Which organs form the central nervous system? (50%)	a) The brain and the spinal cord* b) The cerebrum and the cerebellum c) The cerebrum, the cerebellum and the medulla

	oblongata d) The sensory and motor nerves
23. Which organ produces bile? (23%)	a) Pancreas b) Stomach c) Small intestine d) Liver*
24. What nutritional substances, present in cereals and legumes, provide us with energy? (30%)	a) Lipids b) Proteins c) Carbohydrates* d) Mineral salts
25. Which of the following methods would you use in order to separate a solid from a liquid in a heterogeneous mixture? (78%)	a) Magnetisation b) Distillation c) Filtration* d) Crystallisation
26. What is the bending of light rays called when they pass from a fast medium to a slower one? (63%)	a) Reflexion b) Diffraction c) Attenuation d) Refraction*
27. Which of the following statements about renewable energy sources is correct? (48%)	a) They are considered inexhaustible energy sources* b) They are at present time the most common energy sources c) Thermal solar energy is an energy source that generates electricity by means of photo-voltaic solar panels d) It does not produce any kind of visual impact
28. What kinds of changes modify the composition of matter? (66%)	a) Physical changes b) Chemical changes* c) Biological changes d) None. Matter does not change
29. What characteristic of sound tells us the difference between a high sound and a low sound? (53%)	a) Volume b) Intensity c) Pitch* d) Timber
30. What colour would be obtained if all the colours of the rainbow were mixed? (36%)	a) Black b) White* c) Red d) Yellow