Received: 3 December 2019 | Revised: 25 May 2020 | Accepted: 29 June 2020

DOI: 10.1002/tea.21655

RESEARCH ARTICLE

JRST | WILEY

A curricular Delphi study to improve the science education of secondary school students in Spain

Elena Charro

Department of Experimental Science Education. Faculty of Education, University of Valladolid, Valladolid, Spain

Correspondence
Elena Charro, Department of Experimental Science Education, Faculty of Education, University of Valladolid, 47011 Valladolid, Spain.
Email: elena.charro@uva.es

Abstract

A curricular Delphi study was carried out to help improve and develop scientific training for secondary school students in Spain. Using the Delphi method, the authors analyzed the degree of consensus among more than 100 stakeholders' answers to the question, "What aspects of science education are considered desirable for a (Spanish) citizen?" From their answers, the authors identified a set of five "concepts." Of these, the most relevant was "Science education related to environmental issues and human health, through the use of strategies based on discussion/debate and inquiry-based science learning." This study compares these results with those of previous studies and makes proposals for educational reform based on the implications of its findings.

KEYWORDS

Delphi study, science education, secondary education curriculum

1 | INTRODUCTION

Throughout the twentieth century, the regulation of Spain's education system redirected the focus of science education toward the professional training of future scientists. Recently, however, a growing consensus has developed around the idea that students' scientific competencies must not be constrained to the technological, industrial, and professional applications of the scientific knowledge; that rather, modern citizens of our globalized and increasing technological world should develop basic scientific competencies as part of their education, and integrate these competencies into their day-to-day life and activities (Osborne, Simon, & Collins, 2003;

Schulte, 2015). This opinion is echoed by the European Commission's (2015) push for "science education for responsible citizenship." Spain's compulsory secondary education curriculum strives to inculcate a series of basic competencies in students; among these, competencies in science and technology attempt to help students develop approaches to the physical world through conscious interaction with it (MEC, 2017).

Many authors agree on the importance of improving and enhancing scientific literacy in society in general and in secondary school students in particular. However, it is difficult to distinguish between practical and ideal approaches to and topics for secondary school students' science education, and to determine what students and teachers prioritize and value in science education. Scholars working on these issues often use Program for International Student Assessment (PISA) tests to examine students' performance in key subject areas and study a wide range of educational outcomes, including students' motivation to learn and their perception of their academic and other abilities (Fensham, 2009). They suggest that replacing traditional deductive-inductive study methodologies with inquiry-based methodologies like PISA tests can help pave the road for successful education reform and outcomes (Rocard et al., 2007). In this case, success entails increasing students' interest in and knowledge of science, stimulating students' and teachers' motivation to engage in science education, and reorienting science education toward student acquisition of scientific competencies, which will help them develop into more rational, critical, and active citizens.

In view of the European Union (EU)'s recommendations in its report titled, *Key Competences for Lifelong Learning* (EU, 2006), the present paper sheds light on the need for educational reform and clarifies which areas of scientific knowledge and which teaching methods students perceive to be most relevant to their education. However, the implementation of new methods in science education necessarily implies the coordination of the many actors involved at all stages of education—actors who have, a priori, very different priorities and views of the problem.

The present study took place as part of a European project (PROFILES Consortium, 2010) under the Seventh Framework Program (7FP), which seeks to determine what aspects of science education are most relevant and desirable for citizens of current and future societies. This study seeks evidence of consensus among various actors and stakeholders in Spain's education system regarding which key improvements need to be made to, and which scientific competencies should be emphasized within, Spain's secondary science curriculum. In particular, it seeks evidence of consensus in stakeholders' answers to the following two questions: (1) How would you define a proper level of scientific competence for citizens of present-day society? and (2) What changes need to be implemented in Spain's secondary science curriculum?

This paper presents a descriptive and interpretative curricular Delphi study of science education in Spain. It proceeds as follows. Section 2 briefly describes the Delphi method and performs a review of the relevant literature. Section 3 describes this study's methodology, and Section 4 describes its results. Section 5 discusses the results and performs a critical assessment of the study process itself; this assessment includes a comparison with other Delphi studies on the relationship between science education and contemporary citizenship. Section 5.4 describes the implications of this study's findings. Finally, the article concludes with a series of reflective comments.

2 | THE DELPHI METHOD

The Delphi method is characterized by the use of a battery of questionnaires that are passed successively to a group of participants. It is considered a mixed methodology as it leverages both qualitative and quantitative approaches to data collection.

The Delphi method is characterized by its ability to measure both the subject group's assessment of the problem at hand and the level of the group's consensus on that problem. In short, the Delphi method assumes that a group's most common or representative opinion on a complex question is more significant than any one individual's opinion; therefore, it seeks representative (rather than objective) reflections of group consensus regarding complex problems. Such common, representative answers are extracted by having a fixed panel of representatively expert stakeholders, who belong to different groups surrounding the issue at hand and thus have different views and priorities regarding the issue, take a series of surveys as described below (Linstone & Turoff, 1975; Osborne, Collins, Ratcliffe, Millar, & Duschl, 2003). Thus, the quality and size of this panel is key to the success of the Delphi method (Powell, 2002).

Administering Delphi surveys is an iterative procedure that consists of several rounds. In each round, the panel of stakeholders are given a questionnaire. The questionnaires are adaptive—for example, the design of each round of questionnaires is based on the results of previous rounds. At the end of each round, stakeholders are given the opportunity to modify their answers after the preliminary results have been analyzed by the researchers conducting the study and they have relayed this analysis to the stakeholders. Through this process of giving and incorporating feedback in successive rounds of questioning, stakeholder consensus can be reached and identified (Balasubramanian & Agarwal, 2012). The content of the questionnaires is such that the analysis of their results becomes more quantitative and less qualitative as the process evolves; e.g., questionnaires in the first round may be open questions, and those in the final round may ask stakeholders to rank or rate concepts which have emerged from answers given in previous rounds. The degree of consensus among stakeholders can then be quantified by analyzing the statistical dispersion of the values given to each concept and/or each combination of concepts (Osborne, Collins, et al., 2003).

In short, the Delphi method is a versatile tool that is applied in multiple fields of scholarly inquiry, including economics, business, healthcare, and education. Recent applications of this method in the field of education include Kloser (2014), Ruppert and Duncan (2017), González-García, Blanco-López, España-Ramos, and Franco-Mariscal (2019), and Wan and Bi (2020), among others. Although the Delphi method has some limitations—namely, it is time-consuming and the stakeholders' answers might reflect their own prejudices and preconceived ideas—the fact that it collects data anonymously, that its data are analyzed statistically, and that it is an iterative process are all relative strengths of the Delphi method.

3 | METHODOLOGY

This study sought different experts' reflections on the content and aims of science education in Spain. It collected stakeholders' views and opinions about which aspects of scientific literacy and competency they consider to be relevant and pedagogically desirable for individuals in present-day and near future societies through three rounds of questionnaires, each consisting of five questionnaires.

3.1 | The stakeholder panel

The criterion used to select stakeholder panels in Delphi studies depends on the subject under study. Given the vast number of actors involved in the Spanish education system and their essential and interlocking importance to the education process, we narrowed our panel down to four groups of stakeholders, namely:

- Students (S): secondary school students from various schools and levels.
- Teachers (T): secondary school science teachers.
- Educators (E): university-level science teachers and science teacher trainers.
- Scientists (Sc): researchers who work in various domains of science (e.g., chemistry).

Initially, the stakeholder panel for this study included a total of 127 individuals. These individuals were contacted in the subsequent rounds of the process. The participation of each of the four groups of stakeholders in each round, and what percentage of the total panel each group represented in each round, is compiled in Table 1.

3.2 | Design of this Delphi study

The generic Delphi method procedure is laid out in Figure 1. The five questionnaires (henceforth described as Q1–Q5) were administered both face-to-face and via e-mail. The anonymity of participants was ensured.

In general, Delphi studies begin with a series of open-ended questions (Q1). We separated each round of this study into parts so as to collect, deliver, and incorporate feedback for subsequent rounds of the questionnaire. For example, we administered Q1 in round one, part I of this study, and analyzed the results of Q1 in order to create Q2, which we administered in round one, part II. Round two was also divided into two parts (Q3 and Q4), and from round two we developed Q5, which we administered in round three.

The central research question of this study is: What aspects of science education are desirable and pedagogically meaningful for citizens of present and future societies? The participant question is based on this idea, and it constitutes the open-ended question characteristic of the first round in a Delphi study, but also the central aspect in the second round which is addressed as well in order to attain more differentiated and specific findings and insights. The research questions for round two, part I, were: (a) Which desirable aspects of science education should be prioritized, given the participants' responses? (b) To what extent are these currently realized in practice? and (c) What differences between priorities and practice can be identified in the participants' assessments? The research question for round two, part II, was: What empirically-

TABLE 1 Number of participants per sample group in each round

	Round 1			
Sample group	**Part I**	**Part II**	**Round 2**	**Round 3**
Students	61 (48%)	61 (48%)	27 (32%)	54 (45%)
Teachers	23 (18%)	22 (17%)	20 (24%)	23 (19%)
Educators	22 (17%)	22 (17%)	18 (21%)	21 (18%)
Scientists	21 (17%)	21 (17%)	19 (23%)	21 (18%)
Total	127 (100%)	126 (100%)	84 (100%)	119 (100%)

Note: In parenthesis percentages of composition with respect to the total in the corresponding round.

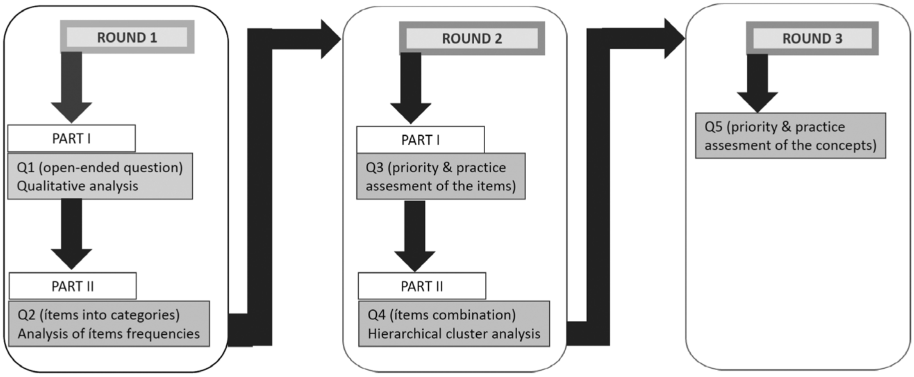

FIGURE 1 Sketch of the different stages of our Delphi study, based on the structure of Bolte (2008)

based conceptions of desirable and contemporary science education can be identified from the participants' statements? The research questions for round three were: (a) Which desirable aspects of science education should be prioritized, given the participants' responses? (b) To what extent are these currently realized in practice? (c) What differences between priorities and practice can be identified in the participants' assessments? and (d) What differences or similarities appear between the four groups' general assessments?

3.3 | Experiment procedure

3.3.1 | Round one

This round involved two steps, including two questionnaires. Q1, administered in round one, part I, read: "What aspects of science education do you consider desirable or important for secondary school students? Please answer with respect to (a) situations and motives that can be taken as a basis or context to facilitate and motivate science education, (b) content and strategies used in the classroom, and (c) knowledge and/or skills which students should learn." Once the stakeholders answered Q1, the researchers generated a set of items from their open-ended responses. They then classified these items into categories and performed a qualitative analysis and generated Q2 using these categories. Q2 provided the stakeholders with these categories and items, and asked a modified question: "What aspects of science education do you consider desirable or important for secondary school students? Choose items (at least one from each category) which you consider the most relevant." The data collected from Q2 were later subjected to statistical analysis.

3.3.2 | Round two

This round saw Q3 and Q4 administered to stakeholders. Q3 sought to explore stakeholders' opinions of which aspects of science education should be prioritized, and the extent to which those aspects of science education are currently being addressed in Spain's education system. Q3 lists the set of items derived from Q1, and consisted of two questions. First, "What degree of relevance should the following items have in science education?” Second, “To what extent are the following items actually being addressed in the current science education system?” Stakeholders were asked to assess their priorities and their sense of how science education was proceeding in practice via a six-tier Likert scale, on which answers were coded from 1 (very low priority/to a very low extent) to 6 (very high relevance/addressed to a very high extent). The collected data were subjected to statistical analysis, and the degree of agreement between the pairs of sample groups' priorities and perceptions was subsequently analyzed via a Mann–Whitney *U* test.

Part II of round two analyzed stakeholders' opinions via Q4, which consisted of the list of items and categories from Q1 and the following question: “Which ideas are the most relevant to secondary school-level science instruction? Express these ideas by creating sets of items, where a set consists of one item from each category.” The data collected from Q4 was subjected to a hierarchical statistical analysis in order to obtain a cluster pattern of the concepts which stakeholders felt were most important and desirable for science education.

3.3.3 | Round three

Q5 was administered in this round. Q5 included the list of concepts from Q1 and the stakeholders two questions: “To what degree should the following concepts be emphasized in science education?” and, “To what extent are the following concepts actually being addressed in the current science education system?” Stakeholders answered Q5 using the same six-tier Likert scale described above. The authors then performed descriptive statistical analyses over the whole stakeholder population and each sample group individually. The pairwise correlations and discrepancies between the results of pairs of groups were later analyzed using the Mann–Whitney *U* test to assess statistically significant differences. The pairwise correlations and discrepancies between results of pairs of concepts were likewise analyzed with the Wilcoxon test. In both cases, the level of statistical significance was set at 0.05 (a 95% confidence interval).

4 | RESULTS

4.1 | Round one

4.1.1 | Round one, part I

Examples of the open-text answers provided by the stakeholders are shown in Figure S1 and S2 (see supplementary online material). The authors read stakeholders' detailed written responses to Q1 and then performed the following qualitative analysis on these responses. First, they used concept-driven coding to label the statements. They created a pre-defined set of codes based on those of Bolte (2008) and the theme of this paper. A total of 88 items were identified through this process. After that, the items were categorized according to a classification system. All codified items were included in the analysis, even if they were mentioned only once. Initially, these items were classified into categories that corresponded to each round of the study—namely, situations and motives (round one/category I), content and strategies (round two/category II), and knowledge and skills (round three/category III). This classification yielded 18, 52, and 18 items, respectively. To classify these items in greater detail, category II was further divided into three more specific categories: contents, fields, and strategies. Category II's initial 52 items were distributed among these sub-categories in the following way: 20 items were classified into contents, 24 items into fields, and 8 items into strategies. This left us with five categories in our final analysis (Table 2). We briefly describe each category below.

- Category I: situations, contexts, and motives

A total of 18 items were assigned to this category. The first three were mainly mentioned by scientists and educators. We assigned the item I.1 labeled "Education/general personal development" to such responses as: "Science education helps students develop their reasoning capacity through the scientific method and contributes to an integral and general formation of individuals' path to becoming an adult" (Educator-1). The item I.3 labeled "Intellectual personality development" was described in the following way: "People in an advanced society such as ours should be interested in general knowledge, not only with respect to history or art, but science as well. Science education contributes to everyone's intellectual development, not only scientists', as many people seem to believe" (Scientist-1). Items I.4 and I.5 were described in the following terms: "Science is taught in the context of the curriculum framework, guaranteeing that the students know the basic concepts of the scientific disciplines, but teachers should connect that content with situations which interest the students" (Teacher-8). Some stakeholders mentioned aspects related to society and media current issues (items I.10 and I.11) for example: "Topics with social and technological interest, …, can be considered good contexts to motivate secondary school students" (Scientist-1) and "Also, current issues shown in different mass media could be good situations to facilitate the science education process …" (Educator-1). Whereas items I.14 to I.16 referred to how students are motivated to learn across different scientific disciplines (biology, chemistry, and physics), item I.17 referred to inter-disciplinary science education. Educators mainly referred to items I.12, I.13, and I.18 in terms of how "non-formal education involving professional conferences and out-of-school experiences such as visiting museums are great activities for motivating students" (Educator-1). Students and scientists described other motivating contexts in items I.6 to I.9, for example, "Natural phenomena such as earthquakes, tsunamis, and hurricanes can be good situations to learn about different scientific subjects" (Scientist-1) were classified under item I.6 ("Nature/natural phenomena").

- Category II: contents and topics

A total of 20 items were assigned to this category. Most of them are related to the basic contents of science education (II.1 to II.6, II.8, II.9, II.12), and were mainly mentioned by scientists and teachers. The relationship of scientific content to everyday life (II.13 to II.20) was usually mentioned by students—for example: "I think that some of the contents of science lessons should be related to everyday life, like those useful for knowing how mobile phones, TVs, and radios work" (Student-5). Scientists' desire that "knowing how the process of scientific inquiry works in different fields be included in secondary school curricula, for example, understanding advances in astronomy" (Scientist-1) were classified under item II.10, "Scientific inquiry." Some stakeholders also considered items II.7 ("Development/growth") and II.11 ("Limits of scientific knowledge") context-appropriate for secondary science education.

- Category III: specific fields

TABLE 2 Classification of the 88 items in five categories as described in the text

Item	I: Situations, context, motives
Code	
I.1	Education/general personal development
I.2	Emotional personality development
I.3	Intellectual personality development
I.4	Students' interests
I.5	Curriculum framework
I.6	Nature/natural phenomena
I.7	Everyday life
I.8	Medicine/health
I.9	Technology
I.10	Media/current issues
I.11	Society/public concerns
I.12	Global references
I.13	Occupation
I.14	Science-biology
I.15	Science-chemistry
I.16	Science-physics
I.17	Science-interdisciplinary
I.18	Out-of-school
Item	**II: Contents and topics**
Code	
II.1	Matter/particle concept
II.2	Structure function/properties)
II.3	Chemical reactions
II.4	Energy
II.5	System
II.6	Interaction
II.7	Development/growth
II.8	Models
II.9	Terminology
II.10	Scientific inquiry
II.11	Limits of scientific knowledge
II.12	Cycle of matter
II.13	Food/nutrition
II.14	Health/medicine
II.15	Matter in everyday life
II.16	Technical devices
II.17	Environment

(Continues)

TABLE 2 (Continued)

II.18	Industrial processes
II.19	Safety and risks
II.20	Occupations/occupational fields
Item	**III: Specific fields**
Code	
III.1	Botany
III.2	Zoology
III.3	Human Biology
III.4	Genetics/molecular biology
III.5	Microbiology
III.6	Evolutionary biology
III.7	Neurobiology
III.8	Ecology
III.9	Inorganic chemistry
III.10	Organic chemistry
III.11	Analytical chemistry
III.12	Biochemistry
III.13	Mechanics
III.14	Electrodynamics
III.15	Thermodynamics
III.16	Atomic/nuclear physics
III.17	Astronomy/space system
III.18	Earth sciences
III.19	Mathematics
III.20	Interdisciplinarity
III.21	Current scientific research
III.22	Consequences of technology development
III.23	History of the science
III.24	Ethics/values
Item	**IV: Knowledge and skills**
Code	
IV.1	(Specialized) knowledge
IV.2	Comprehension/understanding
IV.3	Applying Knowledge/thinking/abstractly
IV.4	Judgement/opinion-forming/reflection
IV.5	Formulating scientific questions/hypotheses
IV.6	Being able to experiment
IV.7	Rational thinking/analyzing/drawing conclusions
IV.8	Working self-dependently/structurally/precisely

TABLE 2 (Continued)

IV.9	Researching/investigating
IV.10	Reading comprehension
IV.11	Communication skills
IV.12	Knowledge about scientific occupations
IV.13	Perception/awareness
IV.14	Sensibility/empathy
IV.15	Social skills/team works
IV.16	Motivation/interest/curiosity
IV.17	Critical questioning
IV.18	Acting reflected and responsibly
Item	**V: Strategies of teaching/learning**
Code	
V.1	Cooperative learning
V.2	Learning in mixed-aged classes
V.3	Interdisciplinary learning
V.4	Inquiry-based science learning
V.5	Learning at stations
V.6	Role play
V.7	Discussion/debate
V.8	Using new media in teaching/learning

A total of 24 items were assigned to this category. The first 18 referred to different scientific disciplines (e.g., botany, zoology, genetics/molecular biology, etc.) mentioned by stakeholders from all four groups. Some of these (e.g., items III.9 to III.16) were only mentioned by scientists. Items III.19 to III.24 were mentioned members of different groups in different contexts—for example, item III.19 ("Mathematics") was described by some teachers in terms of students' need to cultivate their problem-solving skills through mathematics education (Teacher-5). Item III.20 ("Interdisciplinarity") was included by who saw science as "an interdisciplinary subject which can offer motivating contexts to students" (Educator-12). Item III.21 ("Current scientific research") was offered by some students who thought it might be "interesting to know what scientists are working on now" (Student-5). Item III.24 ("Ethics/values") was considered by some scientists to be important for study in some fields, such as genetics.

- Category IV: knowledge and skills

A total of 18 items were assigned to this category, but only seven (items IV.1, IV.3, IV.5, IV.6, IV.7, IV.9, and IV.12) were directly related to science. Some were described by the experts as "To know about the scientific process and to be able to formulate hypotheses, to experiment, to investigate, to analyze, and to construct an argument in relation to scientific knowledge and draw relevant conclusions" (Scientist-1). Stakeholders related the other 11 skills in this category with disciplines outside of science. For example, items IV.2 and IV.10 included the skills necessary to understand information communication, "the ability to express hypotheses, results, and conclusions orally and in writing" (Teacher-6) was included in item IV.11, and items IV.14 and IV.15 included social skills. Other skills were related to more general attitudes, such as item IV.8, "Students' ability to learn autonomously" (Teacher-8). Items IV.4, IV.13, IV.16, IV.17, and IV.18 all had similar meanings and were described in different terms. For example, Educator-1 reflected on item IV.4, "Judgment/opinion-forming/reflection," by stating their belief "that it is necessary to study science since it develops the capacity to judge the truth of some issues that appear in advertisements, on the Internet, and on TV (for example, pseudo-science), and improves one's ability to analyze, reflect, and form opinions."

- Category V: strategies of teaching and learning

A total of eight items were assigned to this category. Only one (V.4, "Inquiry-based science learning") was connected to science. This item was mentioned by some educators as "The best strategy to use, is inquiry-based science learning, given that it is similar to the scientific method, and helps students develop critical thinking skills" (Educator-1). The other seven items in this category are strategies that can be applied in a variety of different subjects. V.1, "Cooperative learning," was mentioned as an important strategy because "Many of jobs nowadays require workers to collaborate with their colleagues" (Scientist-1). Item V.6, "Role play," was mentioned by one teacher as "a good strategy for learning several scientific concepts such as environmental science" (Teacher-1). Item IV.8 "Using new media in teaching/learning" was mentioned because some stakeholders thought YouTube and related platforms could help students "visualize some experiments or understand some concepts" (Student-1).

4.1.2 | Round one, part II

Part II was performed to determine which items identified by the stakeholders in round one, part I they thought were most relevant to science education. This was done by administering Q2 to the stakeholders. Q2 presented stakeholders with the 88 items they had identified via Q1, classified according to the categories laid out above (Table 2). We then performed a quantitative analysis on the data collected from the 126 filled-out Q2s. The number of items extracted from stakeholders from each group, together with the average and median numbers of items per member of each group, is shown in Table S1 (see supplementary online material). Table S1 shows that the average number of items per participant for the total sample (26) is very similar to those of each group sample (ranging from 23 up to 30). Nevertheless, the great discrepancy between the average and median values shows that the distribution of items among the members of any given group is not homogenous.

Q2 also shows us frequency of each item (i.e., the number of times a subcategory or item was selected by each group and by the total of 126 stakeholders). By analyzing the frequencies, we can observe the degree of consensus among the stakeholders. The most frequently mentioned items, by category, are as follows. In category I, "Education/general personal development" (70%), "Nature/natural phenomena" (57%) and "Intellectual personality development" (66%); in category II, "Environment" (48%); in category III, "Human biology" (51%) and "Earth sciences" (49%); in category IV, "Comprehension/understanding" (54%); and in category V, "Using new media in teaching/learning" (73%), "Cooperative learning" (67%), and "Inquiry-based science learning" (64%).

Those items that were mentioned by 50% or more of stakeholders in every sample group represent some degree of consensus among the groups. There were six such items in this study: "Education/general personal development," "Intellectual personality development," "Nature/natural phenomena," "Inquiry-based science learning," "Using new media in teaching/learning," and "Cooperative learning." These constitute the "consensus hexagon" (Figure 2).

4.2 | Round two

4.2.1 | Round two, part I: Descriptive and variance statistical analyses

This round aimed to determine the stakeholders' opinions regarding whether the items from Q1 were being realized in Spain's education system in practice, and the importance they attributed to each item. Q3 was administered in this part of the study. The analysis of the data collected from Q3 took into account stakeholders' separate assessments of priority and practice, as well as the difference between the two: PPD = $X_{\text{Priority}} - X_{\text{Practice}}$ (Bolte, 2008). Figure S3 (see supplementary online material) shows in on spectral way the average priority stakeholders attributed to each of the 88 items. Remarkably, all subcategories ranged above the theoretical mean value (3.5) of the scale, with the only exception of item V.2 with a mean value of 3.43, and thus were considered important by the stakeholders. The item IV.4 "Judgment/opinion-forming/reflection" had the highest mean value (5.2). Item I.1 "Education/general personal development" and item V.4 "Inquiry-based learning" also had high values, 4.89 and 4.60, respectively, in agreement with the results of round one.

Figure S4 (see supplementary online material) shows the average rate at which stakeholders thought each of the 88 items was being addressed in practice. Most of these items had values between 3 and 4 points in the Likert scale—about one point less, on average, than the corresponding priority values in Figure S3. The highest values were given to item II.3 "Chemical reactions" (4.22), item II.14 "Health/medicine" (4.20), and item V.5 "Learning at stations" (4.20). Interestingly, none of them appeared in the first round's consensus hexagon.

We also evaluated the priority-practice difference (PPD) for all 88 items. If an item has a positive PPD value, that implies that stakeholders prioritize it but it is not being addressed in practice. A negative PPD value implies the opposite, and a null PPD value implies complete agreement between priority and practice. Of the 88 items, only item I.5 "Curriculum

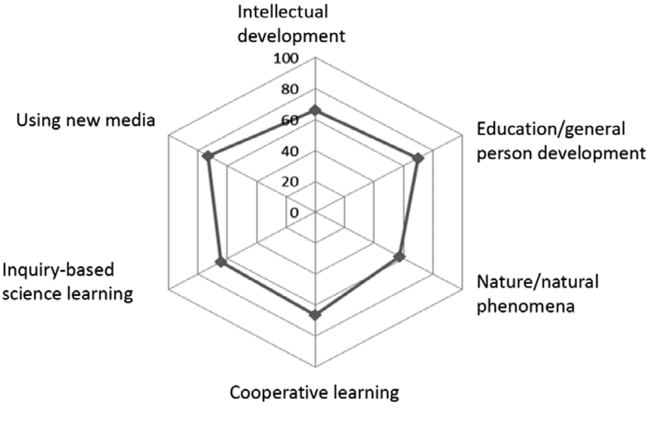

FIGURE 2 Consensus hexagon. The six items of highest overall frequency are located at the vertices of the hexagon. Numbers at the vertices are frequency percentages of the item

framework" had a negative PPD value (−0.3). Item III.20 "Interdisciplinarity," had the highest PPD value (2.63), and item I.14 "Science-biology" had the lowest PPD value (0.08).

The degree of consensus or discrepancy among the stakeholders was analyzed by applying the Mann–Whitney *U* test. In this test, the level of statistical significance was set at 0.05. Table 3 shows the numbers of statistically significant differences between items by the six possible pairs of sample groups: S/T, S/E, S/Sc, T/E, T/Sc, and E/Sc. A total of 29 statistically significant differences were found in the priority assessment (category II had the most of them, with 9), 91 statistically significant differences for the practice evaluation (category II had the most of them, with 30) and 49 statistically significant differences for the PPD. Disagreements between the pairs of sub-sample groups were given percentage values and evaluated alongside the ratio of the number of significant differences by the number of cases (the number of pairs times the number of items). According to this process, the largest degree of consensus corresponds to the lowest percentage values.

Table 4 displays the percentage values of disagreements for all 88 items. For the total sample, the total number of cases was 528 (6 pairs × 88 items). Thus, the percentage of disagreement in priority, practice, and PPD was 5.5, 17.2, and 9.3%, respectively. This reflects consensus rates of approximately 94, 83, and 91% for priority, practice, and PPD, respectively. For each category, discrepancies were evaluated as the ratio of the number of significant differences in that category by the number of cases. The consensus regarding practice is not higher than 87% in any category and regarding priority is not lower than 87%. The best agreement between priority and practice was found in category III, 97 and 87%, respectively.

We also analyzed the consensus between sample groups. For every pair, we calculated the discrepancy as a percentage, considering the number of significant differences in every pair and the total of 88 items. The data are collected and displayed in Table 4. The highest discrepancy for the priority is found between students and teachers (8%), and for the practice is found between students and scientists (27%).

4.2.2 | Round two, part II: Hierarchical cluster analysis

Q4 was administered in this round. It asked stakeholders to make sets of items (one from each category) to express and identify which combinations of concepts they think are important for secondary students' science education.

TABLE 3 Number of statistically significant differences found between the priority (pri) ratings of the items of each category in the six possible pairs of sample groups, and those for practice (pra) (round 2, part I)

Categories	I	II	III	IV	V	Total	I	II	III	IV	V	Total
Pair	Pri	Pri	Pri	Pri	Pri	Pri	Pra	Pra	Pra	Pra	Pra	Pra
S/T	1	3	2	0	1	7	3	1	0	0	0	4
S/E	1	0	0	1	2	4	1	5	2	1	2	11
S/Sc	0	1	0	1	1	3	2	7	7	4	4	24
T/E	1	2	0	3	0	6	1	6	1	3	2	13
T/Sc	1	3	2	0	1	7	6	9	5	7	2	29
E/Sc	0	0	0	1	1	2	2	2	4	2	0	10
Total	4	9	4	6	6	29	15	30	19	17	10	91

TABLE 4 Degree of discrepancy (in parentheses the degree of consensus) in priority evaluation, practice evaluation and priority-practice difference (PPD), for each category (% with respect to the total number of cases) and for each pair of sample groups (% with respect to the total number of items, 88), in (round 2, Part I)

Category (number of cases)	Priority	Practice	PPD
I: Situations, context and motives (108)	4 (96%)	14 (86%)	8 (92%)
II: Contents and topics (120)	8 (92%)	25 (75%)	2 (98%)
III: Specific fields (144)	3 (97%)	13 (87%)	9 (91%)
IV: Knowledge and skills (108)	5 (95%)	16 (84%)	16 (84%)
V: Strategies of teaching/learning (48)	13 (87%)	21 (79%)	17 (83%)
Total number of cases = 528	5.5 (94%)	17.2 (83%)	9.2 (91%)
Group pairs	**Priority**	**Practice**	**PPD**
Students/teachers	8% (92%)	5% (95%)	9% (91%)
Students/educators	5% (95%)	13% (87%)	7% (93%)
Students/scientists	3% (97%)	27% (73%)	26% (74%)
Teachers/educators	6% (94%)	15% (85%)	3% (97%)
Teachers/scientists	7% (93%)	33% (67%)	8% (92%)
Educators/scientists	1% (99%)	11% (89%)	2% (98%)

We then performed a hierarchical cluster analysis on the data collected from Q4 using the Ward method and squared Euclidian distance. We obtained a cluster pattern after gradually summarizing structurally similar responses into first smaller, and then larger clusters (dendrogram). Figure S5 (see supplementary online material) shows the dendrogram obtained with the distribution and allocation of the different items into clusters. Table S2 (see supplementary online material) includes detailed information about each cluster, including the cluster's total number of items (N-items), number of cases (N-cases), and the relative frequency regarding all cases (% cases). There were a total of 857 cases analyzed. By assessing the items collected in every cluster, we defined and described the following five concepts regarding desirable science education:

- Concept A: Science education should include basic scientific concepts, properties, processes and they should be taught via learning by stations and formulating science.
- Concept B: Science education should be geared toward helping students understand social issues and expand and maintain a global frame of reference with regard to these issues, and should focus on helping students develop skills, such as reading comprehension, reflective action, and critical thinking, and might be taught using strategies such as role play.
- Concept C: Science education should relate material and concepts to everyday life, including work, should be taught through the use of cooperative learning strategies, and should focus on helping students acquire competencies that encourage and enable experimentation.
- Concept D: Science education should cultivate individuals' intellectual development, including the acquisition of communication skills, reasoning skills, and the ability to think about problems related to the field, and should be taught using interdisciplinary learning strategies and new media.

- Concept E: Science education should include knowledge related to environmental issues and human health, and should include discussion/debate and inquiry-based learning.

4.3 | Round three

This round aimed to determine the stakeholders' opinions regarding the extent to which concepts A through E were being realized in practice in Spain's education system, and the importance or priority they assigned to each concept. Q5 was administered in this round. The data collected from Q5 was subjected to a descriptive statistical analysis over the whole stakeholder population and over each sample group separately. The results of Round three include several descriptive statistical analyses of stakeholders' assessments of priority and practice, and quantitative measures of PPD. The analysis and descriptions were made on the basis of both the total sample and the four different sample groups. The data were analyzed through SPSS software using descriptive and variance analytical methods. The analysis took both the stakeholders' priority and practice assessments into account individually and then determined the PPD values for each concept. The assessments of the five concepts were tested for statistically significant differences by applying the Wilcoxon signed-rank test. This significance test was applied for the 10 possible pairs of concepts. Statistically significant differences between different sub-sample groups' assessments were identified through the Mann–Whitney *U* test. In both cases, the level of statistical significance was set at 0.05 (a 95% confidence interval).

4.3.1 | Analysis for all stakeholders

First, we analyzed all stakeholders' assessments of concepts A-E were being realized in the current education system and how they ranked the importance of each concept (Figure S6, supplementary online material). The mean value of the Likert scale was 3.5, so the fact that the average priority value of all five concepts was between 4.23 (concept A) and 4.87 (concept E) implies that the stakeholders held all five concepts to be highly important and worthy of priority. The average assessments of how these concepts were being carried out in practice were lower, ranging from 3.29 (concept C) to 3.89 (concept A). The PPD coefficient varied between 1.48 (concepts A and B) and 1.94 (concept C), reflecting a general need to improve how these five concepts are implemented in secondary science education.

Altogether, the stakeholders saw concept E as the most important, and concept A as the one most realized in the current education system. However, the PPD indicates that, at present, all five concepts are practiced less than they are prioritized. The smallest PPD occurs for concept A (0.34), the largest is for E (1.26). The Wilcoxon test shows that three of the 10 pairs of concepts presented statistically significant differences in stakeholders' priority assessments: B/E, C/E, and D/E. Five of the 10 pairs of concepts showed statistically significant differences in stakeholders' assessments of practice: A/B, A/E, B/E, C/E, and D/E (Table S3, supplementary online material).

4.3.2 | Analysis of each group of stakeholders

Next, let us consider each group of stakeholders' assessments separately. Figure 3 shows the mean values of the priority evaluation carried out by each group ranging from 3.6 to 5.2.

The mean values for the five concepts are very close for each sub-sample group, with the exception of the students group. Figure 3 shows also the mean values of each group's evaluation of how the five concepts are realized in practice using a Likert scale which ranged from 1–6, where only the student group gives values ranging from 3.6 and 3.8.

The priorities of all stakeholders as a whole were also reflected in the smaller groups of stakeholders—for example, they do not contradict one another. However, the groups emphasized slightly different things. For example, although students gave the most priority to concept E, they prioritized concept A less than concept B. Among teachers and scientists, however, concept A was prioritized nearly as much or as much as concept E. Students also tended to assess the five concepts as being realized to a higher extent than any of the three other groups of stakeholders did. Likewise, the students perceived the gap between priority and practice to be smaller than any other group of stakeholders did, with the exception of concept E, with a value of 1.24. The largest PPD values occur in scientists for concepts C and D, with 2.24 and 2.09, respectively.

Table S4 (see supplementary online material) displays the number of significant differences found by applying the Wilcoxon test to the 10 pairs of concepts. We can see that the total number of significant differences is higher in practice (12) than in priority (8) assessments'. In all, the student group is the only group that presents discrepancies in the priority assessment and that have no discrepancies in the practice assessment.

Table S5 (see supplementary online material) displays the number of significant differences found using the Mann–Whitney *U* test. A total of 19 significant differences were found

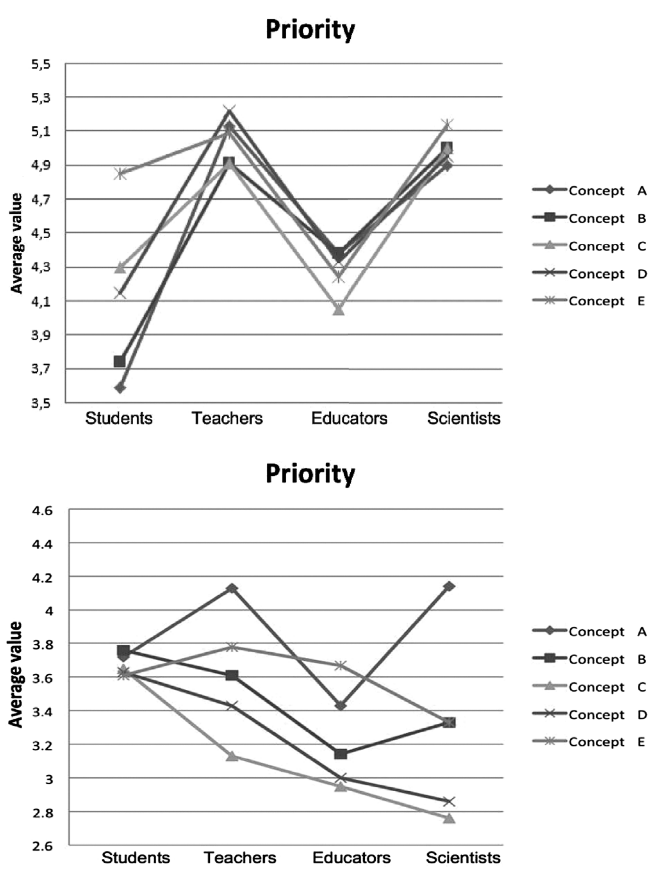

FIGURE 3 Graphical representations of the average values of the Priority and Practice assessment assigned by each sample group to the five concepts

regarding priority assessment pairwise correlations in the same group. The total discrepancy percentage is 63%, evaluated considering 19 of a total of 30 cases (5 concepts × 6 pairs of group comparisons). There were a total of nine differences regarding assessments of practice, which gives a discrepancy of only 30%. No discrepancies (degree of consensus 100%) was found between teachers and scientists for priority and practice assessments.

5 | DISCUSSION AND CONCLUSIONS

This study carried out a curricular Delphi study to determine how various stakeholders in the Spanish education system valued various aspects of science education. It resulted in five concepts regarding science education which should be implemented into the science curriculum at the secondary level. Below, we analyze these findings by comparing them to other, similar Delphi studies, corroborate them via a critical literature review, and discuss the merits, limitations, and implications of this study.

5.1 | Comparison with other Delphi studies

The key concepts extracted from the present study and those of other, similar studies (Blanco-López, España-Ramos, González-García, & Franco-Mariscal, 2015; Kapanadze, Bolte, Schulte, & Slovinsky, 2015; Keinonen, Kukkonen, Schulte, & Bolte, 2014; Ozdem-Yilmaz & Cavas, 2016 and Schulte, 2015) are shown in Table S6 (supplemental online material). Most of the studies displayed in Table S6 identified three key concepts; only the present study and that of Blanco-López et al. (2015) identify five key concepts and follow Osborne, Collins, et al. (2003) methodology. Blanco-López et al. (2015) identify several concepts which are used in this study as well—for example, group work is identified as an important component of science education in both studies. Other studies from Germany (Schulte, 2015) and Turkey (Ozdem-Yilmaz & Cavas, 2016) identified the concept of general science-related education and facilitation of interest in contexts of nature, everyday life and living environment, which is close to concepts C, D, and E in the present study. Keinonen et al. (2014) identified studying scientific knowledge which is needed when moving around and managing in nature, acting according to sustainable development and in societal participation—which is related to concept E in the present study—as the most relevant component of science education. Kapanadze et al. (2015) identified a concept in terms of "General science-related education and facilitation of student's interest in contexts of everyday life using modern and various methods of education" which is close to concept C in the present study.

In addition, the comparatively high number of participants in our study and the fact that they represent stakeholders from many levels of science education make the results of our study more reliable. Unlike other, similar Delphi studies which have been performed with a large overall sample and have seen many participants drop out after the first round (Table 5), our study retained almost 100% of its participants from round to round.

5.2 | Critical view

Many papers have performed similar analyses and suggested the implementation of similar strategies, the development of similar skills and the use of similar topics for science learning

TABLE 5 Comparative of different Delphi studies

Delphi study: Reference	Country	Number of experts round 1	Number of experts round 2	Number of items and categories
Osborne, Collins, et al. (2003)	UK	25	23	30 and no categories
Keinonen et al. (2013)	Finland	187	101	85 into 5 categories
Schulte (2015)	Germany	193	154	88 into 5 categories
Kapanadze et al. (2015)	Georgia	110	97	109 into 5 categories
Ozdem-Yilmaz and Cavas (2016)	Turkey	135	125	157 into 5 parts
Chang Rundgren and Rundgren (2017)	Sweden	100	104	75 into 5 categories
Post, Rannikmäe, and Holbrook (2011)	Estonia	38	85	44 into 4 levels
Blanco-López et al. (2015)	Spain	31	29	40 into 6 dimensions
The present work	Spain	127	84	88 into 5 categories

and teaching. In this study, stakeholders identified concepts D and E as the most relevant to science education, and suggested that science education could be improved by applying teaching strategies centered around discussion, debate, and inquiry in contexts related to environmental issues and human health. These results are corroborated by recent work in the field, which has emphasized the importance of inquiry-based learning for science education, such as Akuma and Callaghan's (2018) systematic review of the literature, which clarified the challenges linked to designing and implementing inquiry-based practical work in secondary school science classrooms. Other examples include Mupira and Ramnarain (2018), who compared the use of inquiry-based strategies with traditional direct didactic approaches and found that the latter produced positive outcomes in conceptual learning and improvements in students' science achievement, and Kang and Keinonen (2018), who analyzed how these strategies and the topics used as contexts for learning affect students' interest and achievement in science. In addition, Ditlevsen, Glerup, Sandøe, and Lassen (2020) found that relating content to current events and human health and environmental issues can generate effective learning contexts for students.

Many publications in the literature have focused on interdisciplinarity, which relates to concept D in the present study. You, Marshall, and Delgado (2017) highlighted the importance of interdisciplinarity in teaching and learning about the carbon cycle, which lies at the heart of climate change and sustainability. Recent work by Sund and Gericke (2020) demonstrates the important contribution of interdisciplinary teaching about environmental issues in secondary school in order to promote the education for sustainable development. Likewise, Yang, Liu, and Gardella Jr.'s (2020) recent work shows how inquiry-based instruction influences students' understanding of interdisciplinary concepts. While standard documents in science education have long promoted interdisciplinary understanding, Spain's education system is presently still oriented toward discipline-based learning.

The importance of discussion, debate, and argumentation in classrooms has been analyzed by Capkinoglu, Yilmaz, and Leblebicioglu (2019) and González-Howard and McNeill (2019). Each of these studies investigated how various teaching techniques promoted argumentation among students and how, through these interactions, students could share ideas with their peers and develop a communal understanding. The importance of acquiring communication skills through science education (as mentioned in concept D) has also been studied extensively (for example, Baram-Tsabari & Lewenstein, 2017; Rakedzon, Segev, Chapnik, Yosef, & Baram-Tsabari, 2017).

The importance of relating science educational content to current events and issues of social concern has been covered by a number of authors. For example, Lindahl, Folkesson, and Zeidler (2019) showed that students demand that content be tied to social issues. Other studies, such as Romine, Sadler, and Kinslow (2017) and Herman, Sadler, Zeidler, and Newton (2018) advocate for this approach to teaching, and Lee, Lee, and Zeidler (2019) found that when teachers use social issues as contexts for science education, students need to make additional efforts to apply what they have learned in the classroom to everyday contexts.

5.3 | Merits and limitations of this curricular Delphi study

The findings of this study indicate that the Delphi methodology is an appropriate way to determine the measure of stakeholders' consensus regarding science education. This work makes two main contributions to the literature: first, it describes how relevant stakeholders see various components of science education to be, and second, it describes the extent to which those components are taught at Spanish schools. These are novel contributions to the relevant literature regarding Spain.

This study has several strong points. First, it takes students' opinions into account—this is important because they are essential stakeholders in education and integrating feedback regarding student motivation is key to the success of any education system. Second, this study had and retained a high number of participants, unlike some other Delphi studies, and so the results are rather reliable. Third, it provides clues and recommendations for how to improve science education. Fourth, it describes the degree to which various stakeholders agree on the relevance of the concepts identified and how well they are implemented in practice. Fifth, it analyzes the differences between stakeholders' assessments of the relevance of these concepts and how they are implemented in practice, which can help us determine which aspects of the existing education system should be modified or improved.

This study also has some limitations. For example, its results are not necessarily generalizable or transferrable to other national contexts. By one hand, the analysis of the results on the extension in the practice of the five concepts are in the secondary schools are only useful for Spain and not to other countries. On the other hand, the opinion about the priority of the concepts are only from Spanish stakeholders and this fact could be considered without application for other countries, however, we consider that scientists and educators probably have a relatively international view on the issue at hand in this study. Finally, we would like to mention that the fact that many publications in the literature corroborated our results despite being conducted in different national contexts leads us to assert that our study should be of interest for international readers.

5.4 | Educational implications

The findings of this study indicate that science education in Spain needs to be reformed. The results might be used to help develop new educational material, teaching models, and secondary teacher training programs. However, the rigidity of the Spanish secondary school curriculum makes such reforms rather difficult, because science is taught in a number of separated science disciplines. Therefore, many socio-scientific issues require involve multiple scientific fields, which means that they can be presented in different disciplines. Physics, Chemistry, Biology, and Geography curricula should strengthen the fusion and contact with each other and break the barriers between disciplines on the basis of highlighting their respective disciplinary peculiarities. Therefore, science curriculum in elementary education should teach students the big ideas and cross-cutting concepts so as to help them to build an overall understanding of the world. In fact, science is a whole and teaching science in separated curricula is only because of the demand of the school education.

The concepts found give the clues to find the main motivational contexts in science education for secondary students: everyday life situations, human health, environmental and technological issues, and social concerns, which can be taking in account for an interdisciplinary learning of the basic concepts of science. Another important clue is the strategy to be used, being inquiry-based science learning proposed by concept E, the most convenient, given that it is similar to the process followed in the scientific method. Science education has to be also the tool for the acquisition of skills, as to be able to do experiments, drawing conclusions or critical-questioning, although some of these skills could be acquired through other non-scientific disciplines, as communication skills. But above all, science education serves not only for having a science-related basic knowledge, it also provides a general education as well as an intellectual and emotional personal development. Thus, aspects as to understand limits of scientific knowledge, or to form an opinion through reflection and rational-thinking, are important to become responsible citizens who take decisions based on scientific arguments.

In both teaching and teacher training this will require the use of different approaches, mainly context-based science teaching and inquiry-based learning, in order to show how scientific knowledge can be integrated with other elements (skills or attitudes) in the context of specific teaching programs designed to develop competencies (Charro, 2017); these approaches would also need to be applied to student assessment (OECD, 2019). Teaching programs of this kind would demonstrate that while competence cannot be achieved without knowledge, the goal is not to possess knowledge but to be able to apply it and solve real-life problems (Lee, 2016).

Given the enormous amount of scientific and technological knowledge that now exists, the task of identifying key aspects is becoming increasingly complex. In this respect, the present analysis would seem to support the literature as regards the need to select a small, common core of knowledge that has both intrinsic and educational value, and which is relevant and applicable to students' everyday lives. Some of the knowledge areas proposed by our panel of experts are included in the concepts, knowledge about the human health and about the environment; they also appear in most of the other publications we have consulted (Wan & Bi, 2020).

The results and conclusions have important implications as regards the role of knowledge within formal science education. One of these derives from the fact that the experts in our study reach a consensus regarding the kinds of knowledge that should be regarded as key aspects of citizens' scientific competence. This suggests that knowledge should cease to be the primary basis of science teaching in schools. This does not mean that scientific knowledge is unimportant, but rather that it should be seen as a necessary but insufficient element for developing the competencies in science that are considered important at different stages of education (Hale, 2013).

Versions of these concepts should form part of the objectives of school science curricula so that all citizens can acquire the skills they require to participate effectively in contemporary society. Indeed, the goal of training responsible citizens who are interested in issues related to science and technology is one that will only be achieved through the concerted input of many sectors of society.

Additionally, the fact that some of the key aspects identified here are not exclusive to science and technology highlights the need to involve other disciplines in the joint development of these competences, which should not be regarded as the sole responsibility of science teachers. Nevertheless, we believe that a formal grounding in science does play a central role in developing the scientific competence of citizens.

During the study the panel of experts offered examples of contexts in which these aspects were important (see Figures S1 and S2). We propose, therefore, that they are transferable to the context of school science teaching. In this regard, we are aware that the five key aspects identified here could be approached within science teaching (Klosterman, Sadler, & Brown, 2012). However, we would argue that the main educational implication of our findings is that these aspects should be treated as an interrelated set of skills. This would imply addressing issues of relevance in the different areas of everyday life (personal, social, professional, etc.) in which students will be required to make decisions, with particular emphasis being placed on those issues over which there is social and/or scientific controversy (Levinson, 2006; Zeidler, Sadler, Simmons, & Howes, 2005).

ACKNOWLEDGMENTS

The author thanks the study participants for their generous cooperation and acknowledges the anonymous referees for their helpful suggestions. This work was supported by the European Union's 7FP.

ORCID

Elena Charro https://orcid.org/0000-0003-2690-2291

REFERENCES

Akuma, F. V., & Callaghan, R. (2018). A systematic review characterizing and clarifying intrinsic teaching challenges linked to inquiry-based practical work. *Journal of Research in Science Teaching*, *56*(5), 619–648.

Balasubramanian, R., & Agarwal, D. (2012). Delphi technique, A review. *International Journal of Public Health Dentistry*, *3*(2), 16–25.

Baram-Tsabari, A., & Lewenstein, B. V. (2017). Science communication training: What are we trying to teach? *International Journal of Science Education, Part B*, *7*(3), 285–300.

Blanco-López, A., España-Ramos, E., González-García, F. J., & Franco-Mariscal, A. J. (2015). Key aspects of scientific competence for citizenship: A Delphi study of the expert community in Spain. *Journal of Research in Science Teaching*, *52*(2), 164–198.

Bolte, C. (2008). A conceptual framework for the enhancement of popularity and relevance of science education for scientific literacy, based on stakeholders' views by means of a curricular Delphi study in chemistry. *Science Education International*, *19*(3), 331–350.

Capkinoglu, E., Yilmaz, S., & Leblebicioglu, G. (2019). Quality of argumentation by seventh-graders in local socioscientific issues. *Journal of Research in Science Teaching*, *57*, 827–855. https://doi.org/10.1002/tea.21609

Chang Rundgren, S. N., & Rundgren, C. J. (2017). What are we aiming for?- a Delphi study on the development of civic scientific literacy in Sweden. *Scandinavian Journal of Educational Research*, *61*(2), 224–239.

Charro, E. (2017). Inquiry based science education strategies for training teachers of secondary schools in environmental issues: A case-study from Spain. *International Journal of Education and Research*, *5*(10), 153–164.

Ditlevsen, K., Glerup, C., Sandøe, P., & Lassen, J. (2020). Synthetic livestock vaccines as risky interference with nature? Lay and expert arguments and understandings of "naturalness". *Public Understanding of Science*, *29* (3), 289–305. https://doi.org/10.1080/00958964.2019.1710444

EU (European Union). (2006). Recommendation 2006/962/EC of the European Parliament and of the Council, of 18 December, on key competences for lifelong learning. Brussels.

European Commission. (2015). *Science education for responsible citizenship*. Luxembourg: Publications Office of the European Union.

Fensham, P. (2009). Real world contexts in PISA science: Implications for context-based science education. *Journal of Research in Science Teaching*, *46*(8), 884–896.

González-García, F. J., Blanco-López, A., España-Ramos, E., & Franco-Mariscal, A. J. (2019). The nature of science and citizenship: A Delphi analysis. *Research in Science Education*, 1–28. https://doi.org/10.1007/s11165-018-9817-5

González-Howard, M., & McNeill, K. L. (2019). Teachers' framing of argumentation goals: Working together to develop individual versus communal understanding. *Journal of Research in Science Teaching*, *56*(6), 821–844.

Hale, E. (2013). From inert knowledge to activated knowledge: Practical ideas for improving student learning. *Philosophy Study*, *3*(4), 312–323.

Herman, B. C., Sadler, T. D., Zeidler, D. L., & Newton, M. H. (2018). A socioscientific issues approach to environmental education). A Socioscientific issues approach to environmental education. In G. Reis & J. Scott (Eds.), *International perspectives on the theory and practice of environmental education: A reader* (pp. 145–161). New York, NY: Springer.

Kang, J., & Keinonen, T. (2018). The effect of student-centered approaches on students' interest and achievement in science: Relevant topic-based, open and guided inquiry-based, and discussion-based approaches. *Research in Science Education*, *48*, 865–885.

Kapanadze, M., Bolte, C., Schulte, T., & Slovinsky, E. (2015). Stakeholders' views on science education—Curricular Delphi study in Georgia. *American Journal of Educational Research*, *3*(7), 897–906.

Keinonen, T., Kukkonen, J., Schulte, T., & Bolte, C. (2013). Stakeholders' views of science education: Finnish PROFILES curricular Delphi study - Second round. *ESERA Conference Proceedings. Strand*, *8*, 160–171.

Keinonen, T., Kukkonen, J., Schulte, T., & Bolte, C. (2014). Stakeholders' Views of Science Education: Finnish PROFILES Curricular Delphi Study – Second Round. In C. P. Constantinou, N. Papadouris, & A. Hadjigeorgiou (Eds.), E-Book Proceedings of the ESERA 2013 Conference: Science Education Research For Evidence-based Teaching and Coherence in Learning. Part 8 (co-ed. M. Ossevoort & J. A. Nielsen) (pp. 160–171). Nicosia, Cyprus: European Science Education Research Association. Retrieved from http://www.esera.org/media/esera2013/Tuula_Keinonen_21Dec2013.pdf (2014-06-12).

Kloser, M. (2014). Identifying a core set of science teaching practices: A Delphi expert panel approach. *Journal of Research in Science Teaching*, *51*(9), 1185–1217.

Klosterman, M. L., Sadler, T. D., & Brown, J. (2012). Science teachers' use of mass media to address socio-scientific and sustainability issues. *Research in Science Education*, *42*(1), 51–74.

Lee, H., Lee, H., & Zeidler, D. L. (2019). Examining tensions in the socioscientific issues classroom: Students' border crossings into a new culture of science. *Journal of Research in Science Teaching*, *57*(5), 672–694.

Lee, Y. C. (2016). Nutrition labelling: Applying biological concepts and reasoning to socio-scientific issues. *School Science Review*, *98*(363), 83–87.

Levinson, R. (2006). Teachers' perceptions of the role of evidence in teaching controversial socio-scientific issues. *Curriculum Journal*, *17*(3), 247–262.

Lindahl, M. G., Folkesson, A. -M., & Zeidler, D. L. (2019). Students' recognition of educational demands in the context of a socioscientific issues curriculum. *Journal of Research in Science Teaching*, *56*, 1155–1182.

Linstone, H. A., & Turoff, M. (1975). The Delphi method: Techniques and applications. In *Reading*. Mass. u.a: Addison-Wesley.

MEC. (2017). Retrieved from: https://www.mecd.gob.es/educacion-mecd/areas-educacion/estudiantes/educacionsecundaria/informacion-general/competencias-basicas.html.

Mupira, P., & Ramnarain, U. (2018). The effect of inquiry-based learning on the achievement goal-orientation of grade 10 physical sciences learners at township schools in South Africa. *Journal of Research in Science Teaching*, *55*(6), 810–825.

OECD. (2019). *PISA 2018 results (Volume I): What students know and can do PISA*. París: OECD Publishing. https://doi.org/10.1787/5f07c754-en

Osborne, J., Collins, S., Ratcliffe, M., Millar, R., & Duschl, R. (2003). What "ideas-about science" should be taught in school science? A Delphi study of the expert community. *Journal of Research in Science Teaching*, *40*(7), 692–720.

Osborne, J., Simon, S., & Collins, S. (2003). Attitudes towards science: A review of the literature and its implications. *International Journal of Science Education*, *25*(9), 1049–1079.

Ozdem-Yilmaz, Y., & Cavas, B. (2016). Pedagogical desirable science education: Views on inquiry-based science education in Turkey. *Journal of Baltic Science Education*, *15*(4), 506–522.

Post, A., Rannikmäe, M., & Holbrook, J. (2011). Stakeholder views on attributes of scientific literacy important for future citizens and employees—A Delphi study. *Science Education International*, *22*(3), 202–217.

Powell, C. (2002). Methodological issues in nursing research. The Delphi technique: myths and realities. *Journal of Advanced Nursing*, *41*(4), 376–382.

PROFILES Consortium. (2010). FP7 negotiation guidance notes—Coordination and support actions. Annex I—Description of work (unpublished).

Rakedzon, T., Segev, E., Chapnik, N., Yosef, R., & Baram-Tsabari, A. (2017). Automatic jargon identifier for scientists engaging with the public and science communication educators. *PLoS One*, *12*(8), e0181742. https://doi.org/10.1371/journal.pone.0181742

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*. Brussels, Belgium: European Commission.

Romine, W. L., Sadler, T. D., & Kinslow, A. T. (2017). Assessment of scientific literacy: Development and validation of the quantitative assessment of socio-scientific reasoning (QuASSR). *Journal of Research in Science Teaching*, *54*(2), 274–295.

Ruppert, J., & Duncan, R. G. (2017). Defining and characterizing ecosystem services for education: A Delphi study. *Journal of Research in Science Teaching*, *54*(6), 737–763.

Schulte, T. (2015). *Desirable science education findings from a curricular Delphi study on scientific literacy in Germany*. Berlin: Springer Spektrum.

Sund, P., & Gericke, N. (2020). Teaching contributions from secondary school subject areas to education for sustainable development—A comparative study of science, social science and language teachers. *Environmental Education Research*, *26*, 772–794. https://doi.org/10.1080/13504622.2020.1754341

Wan, Y., & Bi, H. (2020). What major "socio-scientific topics" should the science curriculum focused on? A Delphi study of the expert community in China. *International Journal of Science and Mathematics Education*, *18*, 61–77.

Yang, Y., Liu, X., & Gardella, J. A., Jr. (2020). Effects of a professional development program on science teacher knowledge and practice, and student understanding of interdisciplinary science concepts. *Journal of Research in Science Teaching*, *57*(7), 1028–1057. https://doi.org/10.1002/tea.21620

You, H. S., Marshall, J. A., & Delgado, C. (2017). Assessing students' disciplinary and interdisciplinary understanding of global carbon cycling. *Journal of Research in Science Teaching*, *55*(3), 377–398.

Zeidler, D. L., Sadler, T. D., Simmons, M. L., & Howes, E. V. (2005). Beyond STS: A research-based framework for socioscientific issues education. *Science Education*, *89*(3), 357–377.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Charro E. A curricular Delphi study to improve the science education of secondary school students in Spain. *J Res Sci Teach*. 2021;58:282–304. https://doi.org/10.1002/tea.21655