Science teacher education on nature of science through explicit and reflective curriculum development

María-Antonia MANASSERO-MAS[1] and Ángel VÁZQUEZ-ALONSO[2,*]

[1] University of the Balearic Islands, Department of Psychology, SPAIN

E-mail: ma.manassero@uib.es

[2] University of the Balearic Islands, Centre of Postgraduate Studies, SPAIN

[*]Corresponding Author's E-mail: angel.vazquez@uib.es

Received 27 Oct., 2020
Revised 23 Jun., 2021

Contents

Abstract

Nature of science (NOS) is a key component of scientific literacy, so science teacher education on NOS is critical to ensure students' scientific literacy. According to mainstream NOS literature, explicit and reflective interventions are recommended to effectively teach NOS. This paper presents an experience for the initial education of secondary science teachers, which focuses on teachers' professional and pedagogical content knowledge development on NOS. A mixed-method approach involving quantitative and qualitative activities was applied to evaluate the effectiveness of the intervention to improve teachers' NOS understanding. The quantitative activity was carried out through a quasi-experimental design with pre-post-test assessments, and the qualitative activity involved the teachers' development of a teaching-learning sequence on a controversial science historical case and their self-reflections on their NOS beliefs, as expressed in the sequence and the answers to the test. The participants are two cohorts of Spanish high-school science teachers enrolled in the Master's course to earn their pre-service education license to teach in Spanish schools. The results diagnosed pre- and post-test teachers' NOS beliefs, their actual improvements, and their self-assessments of their understanding of and changes on NOS. The findings showed an overall modest improvement of teachers' NOS understanding, the differential impact of treatment across cohorts, teachers, and NOS issues, and the difficulty of the self-reflection processes to develop teachers' awareness of their NOS beliefs. Furthermore, this ability was positively and significantly related to post-test and improvement indexes. Finally, the impact, generalization, improvement, and limits of the intervention model are discussed.

Keywords: nature of science; pre-service teacher education; experimental research; intervention assessment; scientific literacy; science competence.

Introduction

The nature of scientific knowledge or nature of science (NOS) represents some interdisciplinary contents of the science curricula about what science is and how science works as a way of knowing and explaining the natural world, which has been elaborated from interdisciplinary perspectives (history, philosophy and sociology of science and technology, and others). The central issue of NOS is knowledge construction and validation, which includes the epistemological principles underlying scientific explanations and the institutional and societal issues involving the relationships and interactions within the community of scientists (internal sociology), the mutual influence between society and the techno-scientific system (the external sociology of science), and the general interactions of science, technology, and society (STS), where education, communication, innovation, scientific policies, socio-scientific issues, etc., emerge. Altogether, these aspects make up NOS (Lederman, 2008; Vesterinen, Manassero-Mas, & Vázquez-Alonso, 2014).

Understanding NOS is a key component of the scientific literacy for all and a perennial goal of science education (Hodson, 2009; Millar, 2006), whose pursuit has developed innovative science curricula to teach NOS contents in many countries (McComas & Olson, 1998; NRC, 1996). The Next Generation Science Standards (NGSS, 2013) recently provided a reinforced, simplified, and renewed curricular vision of NOS by grouping NOS contents into scientific practices and transversal contents. However, some theoretical controversies about NOS conceptualization and the most appropriate NOS issues for curricula still add some complexity (Erduran & Dagher, 2014). Beyond controversies, NOS is widely acknowledged as a meta-cognitive, multifaceted, and dynamic realm with affective components and values related to science practice. Thus, teaching NOS represents a hard innovative challenge due to its meta-knowledge status (i.e., scientific knowledge is tentative) and complex philosophical (coordination of evidence and explanation) and sociological (scientists scrutinize findings) grounds (Adúriz-Bravo, 2014; Matthews, 2012).

The Spanish secondary science curriculum specifies some NOS contents and attainment goals through a block on scientific research, common for all science subjects. However, the overall contents, textbooks, and teaching practices are quite traditional and neglect the teaching of explicit NOS contents in Spanish science education and pre-service teacher education. The innovative challenge for Spanish science education is twofold, teaching NOS issues to all students and educating teachers on NOS. Thus, this study aims to train pre-service Spanish teachers to improve their NOS conceptions and their professional development and to motivate them to teach NOS within the classroom, through an experience with innovative NOS contents about scientific investigations (Manassero-Mas, Bennàssar & Vázquez-Alonso, 2018).

Teachers and Teaching the Nature of science

Innovations usually generate some teachers' resistance to change, which poses a general educational challenge. In the case of NOS teaching, the literature adds two additional hindrances: teachers' uninformed NOS beliefs and lack of teaching resources (Heering & Höttecke, 2014; Donnelly & Argyle, 2011).

For years, research consistently and repeatedly shows that most teachers hold uninformed and poor NOS understanding (Apostolou & Koulaidis, 2010; Bennàssar, Vázquez, Manassero & García-Carmona, 2010; Celik & Bayrakçeken, 2006; Ma, 2009), which is the greatest obstacle to teaching NOS (Irez, 2006; Lederman, 2008; Vázquez & Manassero, 2012). The NOS profile of Spanish secondary science teachers leads to an unsatisfactory competence to teach NOS. Further, the differences between Spanish novices and experienced teachers are non-significant, which means that teaching practice is ineffective for improving teachers' NOS understanding (Vázquez-Alonso, García-Carmona, Manassero-Mas & Bennàssar-Roig, 2013). As it is impossible to teach the contents that teachers do not master, teachers' appropriate understanding of NOS constitutes a necessary condition for quality NOS teaching (Tsai, 2007).

The review of Deng, Chen, Tsai, and Chai (2011) endorses explicit and reflective instruction as the pedagogical conditions for quality and effective NOS teaching, a claim also supported by others (i.e., Lederman, 2008). Explicit teaching refers to the intentional treatment of NOS topics, which involves full educational planning and explicit application in the classroom, and goes beyond indirect instruction (Abd-El-Khalick & Akerson, 2009; Lederman, 2008). Reflection means that learners must perform meta-cognitive and reflective activities about NOS, such as exploration, analysis, discussion, debate, conclusion, argumentation, decision making, etc. (Kucuk, 2008; Deng et al., 2011).

Some NOS teaching resources have currently been developed in different research studies with pre-service teachers. Wong, Hodson, Kwan and Yung (2008) developed real-life context instructional resources in severe acute respiratory syndrome for pre-service teacher education, which explicitly emphasized several aspects of NOS and authentic scientific inquiry (the mutual impact of science and technology, the humanistic character of scientists, and the inseparable links between science and the social environment). Likewise, Adúriz-Bravo and Izquierdo-Aymerich (2009) applied to pre-service science teachers a research-informed instructional resource about the discovery of radium by the Curies. The teachers solved tasks about three NOS ideas (the distinction between discovering and inventing, scientific modeling via abduction, and the hagiographic treatment of Madame Curie) and developed debates about the tasks. The historical scientific controversies have also inspired some educational resources to teach NOS to pre-service teachers. For instance, García-Carmona and Acevedo-Díaz (2017) drew on the Pasteur-Liebig controversy on fermentation to improve teachers' understanding of scientific theories and scientific interpretations; Aragón, Acevedo-Díaz, and García-Carmona (2019) used the Semmelweis case as a didactical tool to impact teachers' ideas on scientific observations, inferences, creativity, methodology, hypothesis, theories, communication, policy, and scientists' personality. Suzuri-Hernandez (2010) elaborated six historical cases, ranging from physics to life sciences, and developed readings and questions to teach the difference between facts and interpretations and controversies among scientists. These and other studies emphasize the effectiveness of the history of science to educate teachers on NOS.

Many studies tested the effectiveness of teaching resources and methodologies to change teachers' NOS views. Kucuk (2008) trained twelve pre-service elementary teachers through an explicit/reflective STS course, and the teachers improved their NOS understandings, except for the relationship and distinction between theories and laws. Bell, Matkins, and Gansneder (2011) researched the impacts of climate change and global warming across two contexts (standalone vs. situated within instruction) and two instructional strategies (implicit vs. explicit) in 75 pre-service elementary teachers. The teachers under explicit instruction made statistically significant gains in their NOS views regardless of the topic context, and the participants under explicit instruction as a stand-alone topic could appropriately apply their NOS understandings to new situations. Tsai (2006) also found that instruction about conceptual change was more helpful than direct instruction to change teachers' views about science. Lotter, Singer and Godley (2009) proved the influence of a secondary science methods program in secondary science pre-service teachers' views and enactment of NOS and inquiry-based instructional practices, through multiple low-stake teaching and reflection experiences (similar to the multiple tasks proposed here). Wan, Wong and Zhan (2013) interviewed twenty-four Chinese science teacher educators about their NOS teaching conceptions, and the educators suggested five key dimensions (value of teaching NOS, NOS content, incorporation of NOS instruction in courses, learning NOS, and the role of the teacher) that provide a useful framework to interpret the practice of teaching NOS. Overall, these studies suggest that the effectiveness of explicit and reflective pedagogy depends somewhat on the specific NOS idea and the teaching context. Finally, Cofré et al. (2019) reviewed the state of the art, concluding that the interventions lasted over one semester for pre-service teachers, that some aspects of NOS (empirical basis, observation and inference, and creativity) are easier to learn than others (tentativeness, theory and law, social and cultural embeddedness, subjective aspects of NOS and "the scientific method"), and that future investigation should focus on the differences between the easy and difficult aspects of NOS and on assessing the effectiveness of different kind of courses.

Rationale

Science teacher education on NOS in Spain is almost inexistent due to the poor definition of NOS contents in science school curricula and the lack of further specific prescriptions on NOS for pre-service teacher education. Hence, this study aims to fill in this gap with the development of an educational experience about scientific investigations, which follows and applies procedures and tools of previous research studies, such as the explicit and reflective framework to improve teachers' NOS beliefs, dispositions, and pedagogical professional development (Abd-El-Khalick, 2012; Bennàssar et al., 2010; Vázquez-Alonso & Manassero-Mas, 2019; Suzuri-Hernandez, 2010; Vázquez, & Manassero, 2013a).

Two general frameworks of learning and teacher development are also taken into account. On the one hand, the explicit goal of improving teachers' NOS beliefs considers learning as a conceptual and metacognitive change (Abd-el-Khalick & Akerson, 2009). On the other hand, teacher learning has recently been established through a global model of teacher professional development, which involves a continuous change in learning, planning, and teaching specific topics (NOS) specifically (explicit and reflective) for a specific purpose (science literacy), which, in turn, means enhancing specific students' NOS understanding at a specific time (Gess-Newsome, 2015; Mesci, 2020). This study joins the "7E learning cycle" structure as a NOS pedagogical model, which deploys the explicit and reflective activities across seven learning milestones: elicit (previous ideas), engage (learners), explore (contents), explain (conclusions), elaborate (results), extend (consequences), and evaluate (Eisenkraft, 2003).

Further, this study aims to research the empirical efficacy of the educational experience to improve teachers' NOS understanding of scientific investigations, as a way towards teacher professional development on NOS. The explicit and reflective framework was developed around some short and specific readings about controversial cases of the history of science that align with previous research (Aragón et al., 2019; García-Carmona & Acevedo-Díaz, 2017; Heering & Höttecke, 2014; Klassen & Froese-Klassen, 2014; Rudge & Howe, 2009). The pre-service teachers' professional development was expected to expand around the elaboration of the readings, the answers to some questions about scientific investigations, the development of a lesson plan, and the self-reflection about their own thinking (Sprod, 2014). The research questions are: What are prospective secondary science teachers' prior NOS beliefs about scientific investigations? How do teachers change their prior NOS beliefs as a result of the experience? Is the experience effective to improve teachers' professional development on NOS?

Methodology

This study applied a mixed-method with a quantitative pre-post-test design and the qualitative analysis of pre-service secondary science teachers' (hereafter "the teachers") productions to answer some open questions about their answers to the test, which corresponds to an embedded design where the qualitative answer data-set provides a supportive role for the primary quantitative test data-set. The second author was the instructor, who delivered the resources, collected teachers' data and productions, and continually assisted the teachers.

Participants

The participants were two cohorts of student teachers (24- to 29-year-olds), enrolled in the compulsory pre-service Master's course to earn their license as secondary teachers. The cohorts included 10 and 9 teachers (6 and 4 women, respectively), who all have a degree in Physics (2, 1), Chemistry (5, 3), Environmental (2, 3), and Health Sciences (1, 2). However, the participants lacked prior teaching or professional experience, any NOS-related activity, even in the Master course, and were blind to the research design so that changes on NOS understanding may be attributable to the treatment.

Instruments

The study applies two research instruments: some didactic resources and the assessment tool. The resources are a set of documents delivered to teachers to scaffold their task, aimed at designing a Teaching Learning Sequence (TLS) on scientific investigations (the explicit global NOS issue for the TLS). The set involves a 7E-based template of the TLS, a short and simple reading on a science controversial historical case (the appendix shows an example), and some exemplary NOS instructional activities on the distinction between facts and explanations and group arguments concerning the reading (Vázquez, & Manassero, 2013b). Teachers were allowed to choose one out of six cases, ranging from Physics to Life Sciences: discovering a new planet, light duality, phlogiston and oxygen, moving continents, spontaneous generation, and pellagra (adapted from Suzuri-Hernandez, 2010). The labels of the lesson plan template are: title, abstract, target students, addressed NOS features, learning prerequisites (if any), school subject and curriculum content (to be infused), purpose and justification, learning goals, timing, methodological orientations, assessment criteria, materials and resources, and learning activities.

Planning lessons is a teachers' basic activity of the Master's course for their professional development. In this case, teachers must self-appropriate and draw inspiration from resources that were new for them, elaborate the learning demands, design and fill in the template, add new resources, make methodology decisions to develop their personal reflections about the resources provided on scientific investigations, and elaborate the TLS. The instructor presented the resources, assigned the tasks to the teachers, and gave continuous feedback and support, aimed at creating a positive disposition toward NOS teaching.

Table I. The ten questions applied as pre-test and post-tests to assess teachers' NOS beliefs

Question number*	Themes / sub-themes	Stem of the question (scenario)	Excerpts of the reading on spontaneous generation that are cues to the question theme
10113	Science as a process	The process of performing science is best described as ...:	This idea, often called "spontaneous generation", is quite understandable. For the next two centuries, the debate on spontaneous generation continued. As more and more observations accumulated against it, people gradually stopped believing in spontaneous generation.
60211	Scientists' characteristics	The best scientist always has an open mind, is impartial and objective in his/her work. These personal characteristics are needed to create better science.	Almost all scientists believed this explanation.
70221	Controversies/ fact-based closure	When a new scientific theory is proposed, scientists must decide whether or not to accept it. Their decision is based objectively on the facts that support the theory; it is not influenced by subjective feelings or personal motivations.	Redi concluded that non-living material does not produce living organisms. To prove this further, he put dead flies and dead worms onto meat inside sealed containers. No living maggots appeared in the containers with either dead flies or dead maggots. Redi was satisfied, but many others did not agree with him. As more and more observations accumulated against it, people gradually stopped believing in spontaneous generation.
70611	Universality of science/ scientists' personality	Having the same basic knowledge, two scientists can develop the same theory independently of each other. The scientists' character does NOT influence the content of a theory.	The Italian scientist Francesco Redi
70621	Universality of science / brilliant scientists	Some brilliant scientists like Einstein have a personal and unique way of seeing things. These creative points of view	… suspected that worms are produced by tiny, invisible eggs, laid by flies on the meat. Other insects, such as butterflies, lay eggs

		determine how other scientists in the same field interpret things.	that become larvae before becoming adults.
90111	Observations and theoretical load	Scientific observations made by competent scientists will be different if they believe in different theories.	People saw that mice and rats suddenly appeared in the barns where the grain had been stored for a while. After waiting for a few days, Redi found that maggots appeared only in the open jars. He also saw how maggots eventually turned into flies.
90411	Fallibility and change	Although scientific investigations are correctly conducted, the knowledge that scientists discover through these investigations may change in the future.	In 1668, … A few centuries ago, people thought that grain could produce mice and rats. For the next two centuries, the debate on spontaneous generation continued.
90611	Scientific method	When scientists investigate, it is said that they follow the scientific method. The scientific method is:	Redi tested his idea by putting pieces of meat into a set of jars. He sealed some of the jars, put gauze over the tops of others, and left others open.
90621	Scientific research / utility	The best scientists are those who follow the steps of the scientific method.	He noticed that meat that is left for several days becomes full of maggots and thought the meat produces the maggots.
90631	Accumulative scientific research	Scientific discoveries occur as a result of a series of investigations, each one supported by the former and leading logically to the next until the discovery is made.	This idea agrees with the religious view that man is made of earth, and with the writings of Aristotle who said that all animals are formed from the four elements - fire, water, air, and earth. As more and more observations accumulated against it, people gradually stopped believing in spontaneous generation.

*The first figure of the question number indicates the NOS dimension; for instance, first figures 1 and 9 of the key number indicate epistemic science themes; first figures 6 and 7 of the key numbers indicate social science themes.

The assessment tool is a standardized multiple-rating ten-item questionnaire that assesses teachers' ideas about ten NOS issues (table I). The ten questions were selected from the pool Views on Science-Technology and Society (VOSTS) according to their fit with the NOS issues involved in the historical case readings (Table I, last column). The VOSTS is an empirically developed item pool with one hundred scenario-based questions; each question's scenario displays a specific NOS issue (represented by a five-digit key number). A list of phrases (beliefs), each representing a reason to explain the NOS issue, follows the scenario (Aikenhead & Ryan, 1992; Manassero, Vázquez & Acevedo, 2001). The respondents rate their agreement (direct score) on each of the phrases within a question (multiple-rating) along a nine-point Likert scale (1 to 9, disagreement to agreement) but the respondents can also leave phrases unanswered to avoid forced choices.

The 9-point rating scale is purposefully broad to gain variance, reduce the problem of unequal ordinal intervals that short scales often display, and allow respondents' refined answers. The sentences convey a range of different views, from informed to uninformed, which were classified in one out of three categories (adequate, plausible, naïve), which mirrors other previous triadic classifications for NOS statements (Abd-El-Khalick & Akerson, 2009; Liang et al., 2009). Thus, to obtain a homogeneous meaning for measurements, a rubric was applied to transform the phrases' direct scores (1-9) into a new index measurement scale within the interval [-1, +1]. The processes of classification and rubric elaboration have been presented and applied elsewhere (Manassero, Vázquez & Acevedo, 2006; Vázquez-Alonso, Aponte, Manassero-Mas & Montesano, 2016; Manassero-Mas, Vázquez-Alonso & Bennàssar-Roig, 2016).

The rubric states a direct-code for adequate phrases (i.e., score 9 recodes as +1), a reverse-code for naïve phrases (i.e., score 1 recodes as +1), and an intermediate-code for plausible phrases (i.e., intermediate score 5 recodes as +1), and so on for the remaining scores, as is habitual in measurement scales. All indexes share a common meaning: the higher (or lower) the index, the better (or worse) informed is the respondent's NOS belief, according to the NOS experts' current views. Thus, high positive indexes indicate informed beliefs (close to experts' views), low negative indexes indicate misinformed beliefs (far from the experts' views), whereas the intermediate index scores around 0 are transitional (Manassero et al., 2001; Vázquez et al., 2006; Vázquez-Alonso et al., 2016).

The new index scale provides a common meaning to all the questions and allows a fine-grained assessment: the set of sentence indexes of each question (NOS issue) provides the individual's profile on the question, and the sentences' mean provides a question index that indicates the respondents' overall NOS understanding of each question (Manassero-Mas et al., 2016). This way, indexes offer high-quality assessment of NOS understanding, rich quantitative descriptions of individual understanding, and an accurate statistical analysis, as suggested by Lederman, Bartos and Lederman (2014, p. 991). Lederman, Wade, and Bell (1998, p. 610) considered VOSTS a valid and reliable instrument for the research of NOS beliefs, and many VOSTS items have been applied to assess teachers' and students' NOS understanding, where acceptable validity and reliability have been documented for large samples (Bennàssar et al., 2010; Vázquez-Alonso et al., 2016).

Finally, teachers' qualitative reflections were independently graded by the researchers through simple rubrics, and solving some slight disagreements by consensus among the researchers.

Procedure

The intervention was designed to provide teachers with some explicit, purposeful, and planned opportunities for learning the NOS-targeted aspects through their own professional development. In the long term, it was aimed at making teachers aware of NOS, inducing favorable dispositions about teaching NOS issues, and encouraging and improving their NOS understanding through the previously adopted theoretical frameworks (learning as change and professional development), centered on teachers' activities and reflections. The intervention was part of the regular assignments of the Master course, and particularly aimed at developing peer co-evaluation and self-evaluation (personal reflective thinking).

1. Initial pre-test assessment. Teachers answered the ten-question assessment tool that establishes the baseline of their prior NOS beliefs about scientific investigations (Table I). The tool triggers the teachers' first personal reflections when providing their reflective answers.
2. Experimental treatment. This stage involved several intertwined part-time activities during three months (instructor's guidance, teachers' personal work, and group work) that deploy diverse opportunities to contact NOS concepts and goals through group discussions and personal activities and to make appropriate connections between prior knowledge and new concepts to teach NOS aspects (NOS pedagogy). The instructor explicitly presented the guidelines, documents, and activities; the Spanish NOS curriculum contents (scientific research) were made explicit, as these contents are mandatory for teachers, and the teachers were required to choose a specific secondary subject and a historical controversial scientific case (Table I) to contextualize their tasks (Manassero-Mas et al., 2018). The instructor related the curriculum and resources and presented the TLS template to the teachers, followed up the TLS elaboration, helped develop teachers' NOS beliefs and their reflective tasks, supervised work progress, organized teachers' class presentations and discussions, and gave continuous feedback and support to focus tasks and create a positive disposition toward NOS teaching (Kattmann & Duit, 1998;

Duit & Treagust, 2003). The teachers developed their lesson plans at their own pace, which enabled them to steadily prompt their reflective thinking (self-appropriation of resources, elaboration of learning demands, completion of TLS template, search for new resources, methodology decision-making, and TLS development), and to formalize and transfer their NOS ideas into their teaching practice. Next, the teachers presented, shared, and discussed their lessons, ideas, and experiences with their classmates (30 minutes), which provided additional reflective opportunities for receiving peers' and instructors' feedback, encouraging additional critical thinking, connections, and personal reflections.

3. Final assessment (post-test). Two weeks after completing the treatment, the ten questions were again answered by the teachers (blind to this design); the initial-final comparison allowed quantifying the impact of the treatment on teachers' NOS understanding across the ten NOS issues (Table I).
4. The instructor elaborated the teachers' initial and final answers to the questions and submitted them to the teachers, who were asked to write a personal reflection about the following topics:
    - Explain the reasons that justify your answers in each question.
    - Compare the initial and final responses to identify changes or continuity in each question and elaborate a reflection explaining and justifying the compared data.

On average, teachers spent about 15 hours of personal work and 5 hours of group presentation and discussion of the lesson plans. Additionally, the second cohort devoted two additional hours of a workshop, as some teachers spontaneously discovered the flowchart "How science works" and asked the instructor for group discussion (http://undsci.berkeley.edu/article/scienceflowchart). As the first cohort lacked this additional reflective workshop, the two cohorts have been treated separately to account for the effect of this factor.

The participants make up a natural convenience group of prospective science teachers (a small group to apply inferential statistics). Thus, the effect size statistic was applied to quantitatively assess the pre-post-test changes, that is, the differences between the mean index scores were computed in standard deviation units. The relevance criterion of the effect size ($d>0.4$) was applied to identify significant changes, whereas the effect size ($-.10<d<.10$) was considered as no change.

Results

The results present the teachers' beliefs across the ten NOS issues and the qualitative productions drawn from their personal reflections throughout the TLS-based intervention. Further, the change of teachers' NOS beliefs was elaborated through pre-test and post-test comparisons and the analysis of the teachers' qualitative reflections.

Teachers' prior beliefs about scientific investigations

The initial individual baseline profiles for the nine teachers of Cohort 1 across the ten assessment questions are represented by the pre-test question mean indexes (Figure 1), which showed some macroscopic trends. First, the question mean scores showed a broad dispersion across teachers, which means that teachers' prior NOS beliefs were quite different. For instance, Question 70611 (Scientists' Personality) presented the largest dispersion across the teachers, and Question 70621 (Brilliant scientists) displayed the lowest dispersion although its top-bottom difference extended over about 0.4 points.

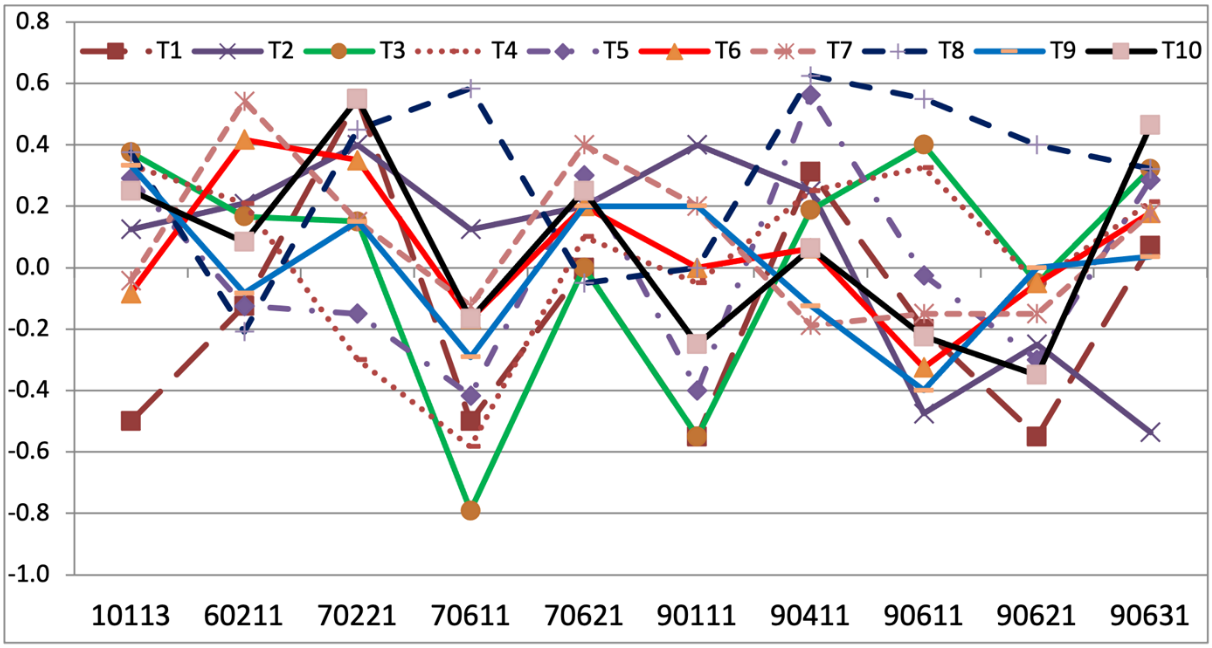

Figure 1. Initial baseline profiles for the ten teachers (Cohort 1) across the ten assessment questions (pre-test mean question indexes)

Second, teachers' prior beliefs depended on the issue; Questions 70221 (Controversy closure) and 90411 (change) displayed the highest and most positive mean scores, whereas the lowest and most negative score corresponded to Questions 70611 (Scientists' Personality) and 90621 (Utility).

Third, a slight global positive trend was perceptible, as Figure 1 displayed more points above zero than below zero across the ten questions; in fact, the cohorts' great mean was slightly positive (+.080).

The above qualitative features of teachers' initial NOS beliefs hold broadly for the second cohort although the representative questions may vary somewhat. Thus, these features allowed concluding that the initial beliefs were very variable across teachers and questions. Moreover, they were not definitely positive, thus confirming previous research that claims that teachers' NOS education is insufficient and inadequate to qualify them for quality NOS teaching.

Change across questions

The appraisal of teachers' change of their NOS beliefs is based on the effect size of the differences between the final (post-test) and initial (pre-test) mean indexes for each of the ten questions (Table II). Overall, teachers improved in most questions, as Table II displays a majority of positive effect sizes (111/190, 58.4%) and fewer negative effect sizes (79/190, 41.6%), revealing an overall improvement in teachers' NOS beliefs.

Table II. Teachers' change profiles across the ten assessment questions between post-test and pre-test mean indexes, represented through the effect size of the differences for each question.

Teachers	**Assessment questions**										**Teacher**
	10113	60211	70221	70611	70621	90111	90411	90611	90621	90631	Mean
T1	0.847	0.236	-0.495	1.025	0.548	0.975	-0.136	0.435	0.802	0.195	0,411
T2	-0.266	-0.428	-0.257	-1.047	0.230	1.109	0.355	1.202	-0.419	1.070	0,194
T3	-0.274	0.094	0.550	0.423	-0.061	-0.305	0.000	-0.139	-0.061	0.049	-0,011
T4	-0.139	0.407	0.108	1.030	-0.493	0.429	0.346	0.141	0.215	0.427	0,226
T5	0.082	0.742	0.336	0.111	-1.520	0.209	-0.965	-0.103	0.668	-0.084	0,016
T6	-0.307	-0.433	-0.567	0.308	-0.921	-1.813	-0.385	0.402	0.525	0.327	-0,209
T7	0.237	-0.551	0.000	0.000	-0.956	-0.500	0.167	0.643	0.394	0.230	0,095
T8	-0.553	-0.113	-0.233	-1.008	1.154	-0.196	-0.474	-0.529	0.000	0.086	-0,193

T9	0.075	0.735	0.259	0.063	0.473	-0.176	0.089	0.442	-0.217	-0.066	0,156
T10	0.506	0.166	-0.807	0.499	0.000	1.746	-0.578	0.859	1.298	-0.093	0,430
T11	1.640	0.000	0.490	0.684	0.757	0.236	0.000	0.331	-0.061	-0.521	0,224
T12	-0.169	0.651	-0.115	0.378	-0.633	-0.704	0.000	0.503	0.000	0.790	0,114
T13	0.554	-0.504	-0.428	0.252	0.132	-0.867	0.000	1.413	-0.316	-0.139	0,115
T14	0.530	-0.214	-0.642	0.000	-0.657	-0.947	-0.150	0.471	0.317	-0.162	-0,052
T15	0.413	0.725	-0.685	0.000	0.000	-0.585	0.521	0.491	-2.079	-0.085	-0,057
T16	-0.306	-0.515	-0.874	0.360	0.000	0.168	0.448	-0.100	-0.279	0.167	-0,085
T17	0.332	-0.855	-0.706	-1.071	-0.510	-0.179	0.317	-0.048	0.182	0.516	-0,141
T18	0.984	-0.577	1.016	-0.208	0.101	-0.817	-0.425	0.511	1.411	-0.139	0,133
T19	-0.244	0.328	-0.226	0.377	0.102	0.000	0.747	0.428	0.000	0.553	0,201
Question Mean	0.208	-0.005	-0.172	0.115	-0.119	-0.117	-0.006	0.387	0.125	0.164	

Five of the ten questions reached a positive mean effect size (d>.10), only three questions reached a negative, albeit nonsignificant, mean effect size (-.20<d<-.10), and the remaining questions hardly changed, with a close-to-zero effect size (.10<d<.10). Question 90611 (Scientific method) reached the highest positive effect size improvement, which is close to the significance threshold, and Question 70221 (Controversy closure) reached the lowest negative effect size improvement.

It is well known that the sample's overall mean decreased the measurement variability, so that it underestimated the effect size. In fact, the effect size of the individual questions across teachers (inner cells, Table II) showed higher scores than its great mean (last row, Table II). For example, Question 90611 (Scientific method), which reached the largest average effect size, showed that 12 teachers exceeded the minimum threshold effect size (d>.40), with some teachers exhibiting high scores (Table II), although the question's great mean did not reach the minimum threshold (d<.40).

Change across teachers

The analysis of the teachers' changes across the final-initial differences (last column, Table II) showed the uniqueness of each teacher's evolution. Most teachers (10) improved their NOS beliefs (d>.10), and two teachers exceeded the minimum significant effect size (d>.40). These two teachers displayed six and five questions that reached a large significant effect size but the presence of one and two questions, respectively, whose effect size was significant and negative, is also noteworthy.

In contrast, only three teachers reached an overall negative effect size (d<-.10), although none of them even reached one half of the effect size threshold. Again, these three teachers with negative global change profiles presented one question that reached a significant and positive effect size. The remaining teachers (6) displayed an overall close-to-zero effect size, and their profiles across questions displayed a variable number of questions with significant effect size scores, either positive or negative.

The teachers' change pattern was quite inhomogeneous and inconsistent across questions, as each teacher simultaneously exhibited significant positive and negative changes, that is, questions that worsened and questions that improved their profile. The teachers with the highest positive improvement profiles also displayed some questions with significant negative effect sizes, and conversely, the teachers with the lowest negative change also displayed some questions with significant positive improvement.

Qualitative Results

The elaboration and presentation of the TLS on investigations and the answers to two self-reflective questions about their test results are some products of the teachers' reflections on NOS issues that could contribute to another qualitative study by themselves. However, the available space limits its presentation to some core results of the first cohort's personal reflections on their question ratings. The instructor provided each teacher with their own ratings and requested two responses about each assessment question: "explain the reasons for your ratings" and "explain the reasons that justify your change/no change in each question when comparing your first responses to the second ones."

Table III. Excerpts of teachers' reflections on the first reflective item (Explain the reasons that justify your ratings) across three significant teachers and three assessment questions

Teacher displaying the highest positive change	Teacher displaying the lowest negative change	Teacher displaying the largest absolute change
10113 I find acceptable those responses that describe the process of doing science as a set of activities-methods	10113 For me, the process of doing science is to discover the mysteries of the universe and the behavior of Nature, using	10113 I agree to a greater or lesser extent with all the answers. Therefore, I have given them all a score above 6. The

designed to explain our environment (nature-world-universe). The phrase D (use of technology to discover Nature) does not seem acceptable because it is limited to a single way of doing science, such as the use of technology.	new technologies and qualitative and quantitative methods.	premises that I think best describe the process of doing science are everything we do, discovering order and proposing and validating explanations, so I gave them a score of 8 (high agreement).
90111 Observations are the results of the experiment directly (without treatment) or sanitized results. Sometimes, treatment can produce incorrect data or erroneous observations.	90111 The observations are always the same but either they cannot explain all the knowledge or they will differ very little, depending on the theories.	90111 Observations must always be objective, but the interpretation of what we see will depend on the mindset and ideas of each person.
90621 I think the best scientists are those who use any method (including the scientific method) to obtain results. The serendipity of many discoveries shows that it is really not necessary to always follow a particular method.	90621 Scientists using the proper scientific method in their field will achieve better results, but chance discovery also produces occasional successes. The scientific method is subject to change but these changes must be based on scientific arguments, to avoid achieving false results.	90621 I agree with the idea that most scientists continue to use this method for their research because it assures valid, clear, logical and accurate results, although these results can also be negative.

Table III presents verbatim some exemplar excerpts of the responses to the first item by three prototypical teachers: the teacher who achieved the overall highest improvement, the teacher who achieved the lowest negative change, and the teacher who achieved the highest absolute changes, either positive or negative. Overall, the excerpts display another facet of the results and exemplify the teachers' diverse evolution patterns across the informed and misinformed NOS conceptions. Further, they also display the subtlety of the nuances that differentiate the teachers' positions from each other, and between informed and misinformed conceptions.

The highest-improvement teacher justifies his responses, considering the centrality of scientific explanations in the process of doing research, stating that observations are much more than simple data and any method is useful to obtain results. The lowest-improvement teacher sustains traditional positivist positions on science (discovering the mysteries of the universe and the behavior of Nature) and observations (observations are always identical) but he considers that theories can project some influence on the observations; he also views the scientific method as the way to achieve better results, while simultaneously acknowledging that "scientific method is subject to change." The highest-change teacher expresses an informed view on observations, a holistic conception of science as a process, and also stresses the centrality of explanations; however, she joins the traditional positivist views in her conception of the scientific method.

The second item requests teachers to identify and explain the changes produced in their thinking when comparing their initial and final responses. Overall, teachers' responses to this requirement followed two different patterns: on the one hand, a concrete pattern, specifying and mentioning the question phrases where they perceived the presence or absence of change; on the other hand, a generic pattern, when the teacher acknowledged change/no change, without mentioning specific evidence.

For instance, Teacher #2 rightly argues for a change in question 70611 (Scientists' personality) through this specific reasoning:

> *"My answers to this question have changed slightly. You can see a trend toward thinking that [a scientist's] character influences the result. Certainly, this change is due to the fact that the debates in class have led me to see the diversity of opinions on the same theories."*

The comment specifies the nature of change and evaluates the amount of change (slight). However, the teacher underestimated the amount of change, as the initial and final scores of most sentences showed quite relevant positive changes, so that the teacher actually achieved a greater change than he perceived.

Globally, teachers' explanations suggest some difficulty to accurately perceive and account for the changes between their initial and final scores. Thus, teachers' self-perceptions of their changes were elaborated by computing hits and misses in each question (Table IV). The procedure compares the teacher's decision expressed in their comments (either change or no change/same) with their actual quantitative ratings (change or no change/same) of the final-initial quantitative differences for each question. If a teacher's position on a question (change or no change) agreed with the actual quantitative difference between question ratings (both change or both no change), the comparison was assessed as a hit. On the contrary, if the two traits were different (change vs. same), a miss was assigned.

Table IV. Frequency of teachers' hits and misses drawn from the reflections elicited by the second item (Explain the reasons that justify your changes) when teachers compared their quantitative ratings on the assessment questions

	Hits				**Misses**			
	Teacher/Ratings		**Teacher/Ratings**		**Teacher/ Ratings**		**Teacher/Ratings**	
	change/change		**same/same**		**change/same**		**same/change**	
	Cohort 1	**Cohort 2**	**Cohort 1**	**Cohort 2**	**Cohort 1**	**Cohort 2**	**Cohort 1**	**Cohort 2**
Teacher 1	1	5	4	1	2	2	3	2
Teacher 2	4	5	4	1		2	2	2
Teacher 3	3	3	4	4	3			3
Teacher 4	4	3	3	4	3	1		2
Teacher 5	2	5	2	1	6	1		3
Teacher 6	4	5	3		2	2	1	3
Teacher 7	6	2	1	4	2	4	1	
Teacher 8	2	3	4	5	2	1	2	1
Teacher 9	4	4	1	4	1	1	4	1
Teacher 10		6		1		1		2

The most interesting results of Table IV are not the perception of hits (left side), as the teachers were expected to be able to correctly identify their changes, but the perception of errors (misses), because they were unexpected (right side), taking into account that the evidence (teachers' ratings) was provided to the teachers. Overall, the teachers made about 34% of mistakes when they assessed their conception changes. Furthermore, there were still two types of misses: on the one hand, when the teacher perceived that their rating had changed and the evidence of ratings showed there was no change (change/same); on the other hand, when they did not perceive change, and the ratings showed that the change was relevant (same/change). Both errors were split nearly into halves, although the former (change/same) was more frequent.

The procedure allows the identification of the teachers according to their self-perceptions of changes; Teacher #5 displayed the highest number of errors (misses) about his perception of changes (6 change/same errors), and in contrast, Teachers #8 and #9 (cohort 2) displayed the lowest number of errors (2).

All in all, the teachers' qualitative reflections contributed to confirming many of the quantitative assessment results drawn from the test indexes. Further, teachers' reflections also showed the many mistakes they make when trying to identify changes in their NOS learning and beliefs through self-assessment, thus supporting the need to deepen the reflections on NOS beliefs.

Relationship between qualitative and quantitative results

The main quantitative products are the indexes of the questionnaires, which represent the teachers' initial and final responses to the assessment instrument questions. The main qualitative products of this experience are the TLS and the reflections on two items (self-justification of scores and self-perception of the changes between the initial and final assessment). The qualitative products of the education course were graded by the researchers through rubrics, with consensus-building in case of disagreement on a rubric cell (prior to grading).

The relationships between the quantitative products, questionnaire indexes, and the qualitative products, quantified through the grades obtained by the teachers, were analyzed through Pearson correlations (Table V). For each teacher, the analysis considered the following variables: initial and final assessment average indexes, the average test change scores (quantitative variables), and the average of the three grades of the qualitative products (TLS, and the answers to the two items about the analysis and identification of changes).

Table V. Pearson correlation coefficients between the test indexes and the grade of the reflective identification of the improvement task for the two cohort teachers (n = 19)

	Change identification (grade)	
	Pearson (r)	**p-significance (two-tailed)**
Pre-Test Scores	.071	.773
Post-Test Scores	.568	.011
Improvement	.586	.008

Most variables showed irrelevant correlation coefficients that lack statistical significance except for the qualitative variable change identification, which significantly correlated with the final evaluation indexes (post-test scores) and the NOS understanding improvement indexes (table V). These results mean that the higher the teachers' effectiveness to identify changes in their NOS beliefs, the greater their index in the final evaluation (post-test), and the higher the improvement in their NOS understanding achieved through the education course. It should also be noted that the TLS score and the justification reflection score did not significantly correlate with the test indexes.

Discussion and Conclusions

A convenience sample of prospective secondary science teachers enrolled in their Master's degree education were explicitly and reflectively taught a concrete and specific NOS theme (investigations) through a mixed-method, instead of teaching several NOS themes or a very extensive one, as has been usual in other studies (Hanuscin, Lee & Akerson, 2011; Lederman, 2008; Tsai, 2007). The process included several curriculum and professional development activities (design, planning, creation, presentation, and discussion of teaching-learning sequences), which are a part of basic teacher education and involve group discussions and personal self-reflection on investigations. The impact of these activities on teachers' understanding of investigations was assessed by a quantitative tool and the qualitative analysis of the reflective activities (Heering & Höttecke, 2014).

The quantitative assessments show that teachers' improvements on NOS understanding are clear, although modest and heterogeneous. Overall, most of the 190 assessment indicators of change (ten questions for nineteen teachers) are positive and represent an improvement, most of the teachers (10) reach a positive effect size mean, against a minority (3) who obtained a negative effect size mean. The results for the 10 assessment questions present a similar pattern, as most (5) achieve a positive effect size, a minority (3) achieve a negative effect size, and two have a zero effect size. These findings add to the growing body of knowledge that empirically supports the efficacy of explicit and reflective teaching methods for improving teachers' NOS understanding (Akerson, Abd-el-Khalick & Lederman, 2000; Bell et al., 2011; Lederman, 2008).

However, when analyzing the complex heterogeneity of the teachers' changes across questions and vice-versa, the study contributes a nuanced and detailed model of changes. The heterogeneity means that teachers' changes simultaneously include some questions that improve and some that worsen their profile. In particular, the highest-progress teachers display some questions with a relevant negative effect size, and the lowest-progress teachers show some questions with a relevant positive effect size. Similarly, the questions achieving the highest mean positive changes involve teachers with high negative changes and the questions with the most negative changes involve teachers with high positive changes. The main consequence of these heterogeneous changes suggests that all the teachers partially improve their NOS beliefs but the improvement depends largely on the teacher and the question. Another finding concerning the change heterogeneity indicates that the assessment method can personalize teachers' NOS progress across strengths and weaknesses (Cook & Buck 2013), which are different for each NOS issue (Cofré et al., 2019).

Another contribution is the significant relationship between teachers' effectiveness to identify their changes on NOS beliefs (reflectivity) with the final and improvement indexes, which show that teachers with the highest self-reflection ability are probably the teachers with the highest improvement rates of NOS understanding. This finding suggests that developing some of the cognitive abilities involved in self-reflection (such as critical thinking and metacognition) may lead to greater progress on NOS understanding (Vázquez-Alonso & Manassero-Mas, 2018).

Altogether, the former findings should be interpreted within the context of the short time and modest experience carried out. First, the quantitative interpretation of improvements is limited to the comparisons that allow the data quantification scale, which is open to new analysis and different interpretations. Second, the intervention does not contain lecturing about NOS, and its time is shorter than in other studies, which usually last several months (Cofré et al., 2019; Deng et al., 2011) so that the modest effects of such a small intervention can be proportionally considered. Third, teachers' changes suggest that the impact of the intervention is mixed and differential, as it seems to be heterogeneously endorsing some misinformed beliefs and weakening some adequate beliefs, instead of homogeneously promoting informed NOS beliefs (Cofré et al., 2019; Tsai, 2006). This irregular impact of treatments is also a challenge to improve its effectiveness, through explicitly reinforcing the informed ideas and disproving the misinformed ones through discussions (Abd-El-Khalick, 2012).

Another prospective contribution is the model developed for teaching NOS to pre-service science teachers, based on their professional development through NOS pedagogical competences such as planning NOS lessons, involving self-reflection and feedback from others, similar to Celik and Bayrakçeken (2006). The standardized model of tasks and the quantitative assessment of changes help to apply the model without instructors' deep knowledge of NOS (Bennàssar et al., 2010; Manassero et al., 2016).

The teachers' qualitative reflection on their quantitative responses confirms the validity of teachers' answers and deepens our understanding of the complexity of teachers' thinking about NOS (Abd-El-Khalick, 2012; Hanuscin et al., 2011). In addition, the quantitative assessment contributes to reducing the risk of idiosyncrasy, facilitates the inferential statistical analyses of teachers' views, which is difficult with qualitative assessments (Brunner, Summers, Myers & Abd-El-Khalick, 2016), and allows standardized comparisons of experiences and researches (Abd-El-Khalick, 2012).

Some limitations of this study relate to the convenience sample of the participants and the short time of the intervention. We therefore recommend further studies with larger representative samples and deeper interventions. Moreover, the relationships between the lesson design and teachers' NOS beliefs, and how the lesson development supports NOS understanding would be also worth studying. Finally, another limitation of the study, and simultaneously an opportunity for improvement, is methodological; the use of the questions' average scores to represent teachers' NOS conceptions is a limitation, as it is well known that the average statistic has a centralizing effect that sharply reduces the variability of the averaged data (in this case, the sentence indexes within each question). The opportunity stems from shifting the analysis of the pre-post differences from question average to sentence indexes, which conserve the entire data variability.

In sum, the explicit-reflective NOS model for teacher education presented herein contributes to the field through some findings: heterogeneous improvement of science teachers' understanding of NOS through reflection embedded in the professional development practices (pedagogy and design), standardized questionnaire-based assessment, quantitative and qualitative data and results, and inferential diagnosis of teachers' weaknesses and strengths on NOS. Furthermore, the modest changes suggest proposals for improvement: making the personal reflection processes longer and deeper, deepening the specific pedagogical content knowledge of NOS (Burton, 2013; Mesci, 2020), and adding lecturing time to actively guide informed and misinformed NOS beliefs (Akerson, Morrison & Mcduffie, 2006).

Acknowledgements

Project EDU2015-64642-R funded by the Spanish State Research Agency and the European Regional Development Fund, European Union. On behalf of all the authors, the corresponding author states that there is no conflict of interest.

References

Abd-El-Khalick, F. (2012). Nature of Science in science education: Toward a coherent framework for synergistic research and development. In B. J. Fraser et al. (Eds.), *Second international handbook of science education* (pp. 1041-1060). New York: Springer. http://dx.doi.org/10.1007/978-1-4020-9041-7_64.

Abd-el-Khalick, F. & Akerson, V. (2009). The influence of metacognitive training on preservice elementary teachers' conceptions of nature of science. *International Journal of Science Education, 31*, 2161-2184.

Adúriz-Bravo, A. (2014). Revisiting school scientific argumentation from the perspective of the history and philosophy of science. In M. R. Matthews (Ed.), *International handbook of research in history, philosophy and science teaching* (pp. 1443-1472). Dordrecht: Springer.

Adúriz-Bravo, A. & Izquierdo-Aymerich, M. (2009). A research-informed instructional unit to teach the nature of science to pre-service science teachers. *Science & Education, 18*, 1177-1192.

Aikenhead, G. S. & Ryan, A. G. (1992). The development of a new instrument: "Views on science-technology-society" (VOSTS). *Science Education, 76*, 477-491.

Akerson, V. L., Abd-el-Khalick, F. & Lederman, N. G. (2000). Influence of a reflective explicit activity-based approach on elementary teachers' conceptions of nature of science. *Journal of Research in Science Teaching*, 37(4), 295-317.

Akerson, V. L., Morrison, J. A. & Mcduffie, A. R. (2006). One course is not enough: Preservice elementary teachers' retention of improved views of nature of science. *Journal of Research in Science Teaching, 43*, 194-213.

Apostolou, A. & Koulaidis, V. (2010). Epistemology and science education: A study of epistemological views of teachers. *Research in Science & Technological Education, 28*, 149–166.

Aragón, M. M., Acevedo-Díaz J. A. & García-Carmona, A. (2019). Prospective biology teachers' understanding of the nature of science through an analysis of the historical case of Semmelweis and childbed fever. *Cultural Studies of Science Education, 14*(3), 525-555. http://dx.doi.org/10.1007/s11422-018-9868-y

Bell, R., Matkins, J. J. & Gansneder, B. M. (2011). Impacts of contextual and explicit instruction on preservice elementary teachers' understandings of the nature of science. *Journal of Research in Science Teaching, 48*, 414–436.

Bennàssar, A., Vázquez, A., Manassero, M. A., & García-Carmona, A. (Coor.) (2010). *Science, technology and society in Latin America: An assessment of the understanding of the nature of science and technology*. Madrid: OEI. www.oei.es/salactsi/DOCUMENTO5vf.pdf.

Brunner, J., Summers, R., Myers, J. Y. & Abd-El-Khalick, F. (2016). Toward quantifying responses to the views of nature of science questionnaire: Empirically investigating qualitative coding. *NARST 2016 Annual International Conference Abstracts* (pp. 126-127). https://www.narst.org/annualconference/2016conference.cfm

Burton, E. P. (2013). Student work products as a teaching tool for nature of science pedagogical knowledge: A professional development project with in-service secondary science teachers. *Teaching and Teacher Education, 29*, 156–166.

Celik, S. & Bayrakçeken, S. (2006). The effect of a 'Science, Technology and Society' course on prospective teachers' conceptions of the nature of science. *Research in Science & Technological Education, 24*, 255–273

Cofré, H., Nuñez, P., Santibáñez, D., Pavez, J. M., Valencia, M., & Vergara, C. (2019). A critical review of students' and teachers' understandings of nature of science. *Science & Education, 28*, 205-248.

Cook, K. L. & Buck, G. A. (2013). Pre-service teachers' understanding of the nature of science through socio-scientific inquiry. *Electronic Journal of Science Education*, 17. http://ejse.southwestern.edu

Deng, F., Chen, D.-T., Tsai, C-C & Chai, C.-S. (2011). Students' views of the nature of science: A critical review of research. *Science Education*, *95*, 961–999.

Donnelly, L. A. & Argyle, S. (2011). Teachers' willingness to adopt nature of science activities following a physical science professional development. *Journal of Science Teacher Education*, *22*, 437-457.

Duit, R. & Treagust, D. (2003). Conceptual change: A powerful framework for improving science teaching and learning. *International Journal of Science Education*, *25*, 671–688.

Eisenkraft, A. (2003). Expanding the 5E Model. *The Science Teacher*, *70*(6), 56-59.

Erduran, S. & Dagher, Z. R. (2014). *Reconceptualizing the nature of science for science education, scientific knowledge, practices and other family categories*. Dordrecht: Springer.

García-Carmona, A. & Acevedo-Díaz J. A. (2017). Understanding the Nature of Science Through a Critical and Reflective Analysis of the Controversy Between Pasteur and Liebig on Fermentation. *Science & Education*, 26, 65–9.

Gess-Newsome, J. (2015). A model of teacher professional knowledge and skill including PCK: results of the thinking from the PCK Summit. In A. Berry, P. Friedrichsen & J. Loughran (Eds.), *Re-examining pedagogical content knowledge in science education*. New York: Routledge. https://doi.org/10.4324/9781315735665.

Hanuscin, D. L., Lee M. H. & Akerson V. L. (2011). Elementary teachers' pedagogical content knowledge for teaching the nature of science. *Science Education*, *95*, 145–167.

Heering, P. & Höttecke, D. (2014). Historical-investigative Approaches in Science Teaching. In M. R. Matthews (ed.), International handbook of research in history, philosophy and science teaching (pp. 1473-1502). Dordrecht: Springer.

Hodson, D. (2009). *Teaching and learning about science: Language, theories, methods, history, traditions and value*. Rotterdam, Holland: Sense Publishers.

Irez, S. (2006). Are we prepared? An assessment of preservice science teacher educators' beliefs about nature of science. *Science Education*, *90*, 1113-1143.

Kattmann, U. & Duit, R. G. H. (1998). The model of educational reconstruction. In H. Bayrhuber & F. Brinkman (Eds.), *What-Why-How? Research in didaktik of biology* (pp. 253–262). Kiel, Germany: IPN.

Klassen, S. & Froese-Klassen, C. (2014). Science Teaching and Narratives: Theory and Methodology. In M. R. Matthews (Ed.), *International handbook of research in history, philosophy and science teaching* (pp. 1503-1529). Dordrecht: Springer.

Kucuk, M. (2008). Improving Preservice Elementary Teachers' Views of the Nature of Science Using Explicit-Reflective Teaching in a Science, Technology and Society Course. *Australian Journal of Teacher Education*, *33*. http://dx.doi.org/10.14221/ajte.2008v33n2.1

Lederman, N. G. (2008). Nature of science: Past, present, and future. In S. K. Abell & N. G. Lederman (Eds.), *Handbook of research on science education* (pp. 831-879). Mahwah, NJ: Erlbaum.

Lederman, N. G., Bartos, S. A. & Lederman, J. S. (2014). The development, use, and interpretation of nature of science assessments. In M. R. Matthews (Ed.), *International handbook of research in history, philosophy and science teaching* (V. 2, Ch. 29, pp. 971-998). Dordrecht: Springer.

Lederman, N. G., Wade, P. D. & Bell, R. L. (1998). Assessing understanding of the nature of science: A historical perspective. In W. F. McComas (Ed.), *The nature of science in science education: Rationales and strategies* (pp. 331-350). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Liang, L. L., Chen, S., Chen, X., Kaya, O. N., Adams, A. D., Macklin, M. & Ebenezer, J. (2009). Pre-service teachers' views about nature of scientific knowledge development: an international collaborative study. *International Journal of Science and Mathematics Education 7*, 987-1012.

Lotter, C. & Singer, J. & Godley, J. (2009). The influence of repeated teaching and reflection on preservice teachers' views of inquiry and nature of science. *Journal of Science Teacher Education*, *20*, 553-582.

Ma, H. (2009). Chinese secondary school science teachers' understanding of the nature of science emerging from their views of nature. *Research in Science Education, 39*, 701-724.

Manassero-Mas, M. A., Bennàssar Roig, A. J., & Vázquez-Alonso, A. (2018). Learning to think and act as scientists: the role of cooperative games. In L. Gómez Chova, A. López Martínez, I. Candel Torres (Eds.), EDULEARN18 Proceedings, 10th International Conference on Education and New Learning Technologies, (pp. 2181-2189). Valencia: IATED Academy. https://library.iated.org/publications/EDULEARN18

Manassero, M. A., Vázquez, A., & Acevedo, J. A. (2001). *Avaluació dels temes de ciència, tecnologia i societat* [Assessment of the topics of science, technology and society]. Palma de Mallorca: Conselleria d'Educació i Cultura.

Manassero-Mas, M. A., Vázquez-Alonso, A., & Bennàssar-Roig A. (2016). Effectiveness of teaching the nature of science: global assessment from students' responses. In J. Lavonen, K. Juuti, J. Lampiselkä, A. Uitto, & K. Hahl (Eds.), *Electronic Proceedings of the ESERA 2015 Conference. Science education research: Engaging learners for a sustainable future*, Part 6. M. Izquierdo & V-M. Vesterinen (co-eds.), (pp. 828-835). Helsinki, Finland: University of Helsinki. http://www.esera.org/media/eBook%202015/eBook_Part_6_links.pdf

Matthews, M. R. (2012). Changing the focus: From nature of science (NOS) to features of science (FOS). In M. S. Khine (Ed.), *Advances in nature of science research. Concepts and methodologies* (pp. 3–26). Dordrecht, Holland: Springer.

McComas, W. F. & Olson, J. K. (1998). The nature of science in international science education standards documents. In W. F. McComas (Ed.), *The nature of science in science education: Rationales and strategies* (pp. 41-52). Dordrecht, Holland: Kluwer Academic Publishers.

Mesci, G. (2020). The Influence of PCK-Based NOS Teaching on Pre-service Science Teachers' NOS Views. *Science & Education*, **29,** 743–769. http://dx.doi.org/10.1007/s11191-020-00117-7

Millar, R. (2006). Twenty-first century science: Insights from the design and implementation of a scientific literacy approach in school science. *International Journal of Science Education*, *28*, 1499-1521.

NGSS Lead States (2013). *Next Generation Science Standards: For States, By State*s. Washington, DC: The National Academies Press.

NRC National Research Council. (1996). *National science education standards*. Washington, DC: The National Academies Press.

Rudge, D. W. & Howe, E. M. (2009). An explicit and reflective approach to the use of history to promote understanding of the nature of science. *Science & Education*, 18, 561–580.

Sprod, T. (2014). Philosophical inquiry and critical thinking in science education. In M. R. Matthews (ed.), *International handbook of research in history, philosophy and science teaching* (pp. 1531-1564). Dordrecht: Springer.

Suzuri-Hernandez, L. J. (2010). *Exploring school students' views of the nature of science*. PhD Thesis University of York, Department of Education.

Tsai, C-C. (2006). Reinterpreting and reconstructing science: Teachers' view changes toward the nature of science by courses of science education. *Teaching and Teacher Education, 22*, 363–375.

Tsai, C-C. (2007). Teachers' scientific epistemological views: The coherence with instruction and students' views. *Science Education, 91*, 222-243.

Vázquez-Alonso, Á., Aponte, A., Manassero-Mas, M-A. & Montesano, M. (2016). A teaching–learning sequence on a socio-scientific issue: analysis and evaluation of its implementation in the classroom. *International Journal of Science Education, 38*(11), 1727-1746. http://dx.doi.org/10.1080/09500693.2016.1212289

Vázquez-Alonso, Á., García-Carmona, A., Manassero-Mas, M. A. & Bennàssar-Roig, A. (2013). Science Teachers' Thinking About the Nature of Science: A New Methodological Approach to Its Assessment. *Research in Science Education*, 43, 781–808. http://dx.doi.org/10.1007/s11165-012-9291-4

Vázquez, Á., & Manassero, M. A. (2012). The selection of content to teach the nature of science and technology (part 1): A review of the contributions of didactic research. *Revista Eureka sobre Enseñanza y Divulgación de las Ciencias, 9*(1), 2-33.

Vázquez, Á., & Manassero, M. A. (2013a). Understanding an aspect of the nature of science and technology: An innovative experience for preservice teachers. *Revista Eureka sobre Enseñanza y Divulgación de las Ciencias*, 10, 630-648.

Vázquez, Á., & Manassero, M. A. (2013b). Historical resources for teaching about scientific decision-making. First Autumn School on History of Science and Education, 14-16 November, Barcelona (Spain). http://schct.iec.cat/Web1AutumnSchool/FirstAutumnSchool.html.

Vázquez-Alonso, Á. & Manassero-Mas, M. A. (2018). Beyond scientific understanding: science education to develop thinking. *Revista Electrónica de Enseñanza de las Ciencias*, 17(2), 309-336.

Vázquez-Alonso, A. & Manassero-Mas, M. A. (2019). A serious game to teach epistemic knowledge: the pleasure of learning to think and act like a scientist. In O. Levrini & G. Tasquier, (Eds.), *Electronic Proceedings of the ESERA 2019 Conference. The beauty and pleasure of understanding: engaging with contemporary challenges through science education*, Part 16, F. Corni & A. Spyrtou (co-ed.), (pp. 1884-1893). Bologna: ALMA MATER STUDIORUM – University of Bologna. https://www.esera.org/publications/esera-conference-proceedings/esera-2019

Vázquez, A., Manassero, M. A., & Acevedo, J. A. (2006). An analysis of complex multiple-choice science-technology-society items: Methodological development and preliminary results. *Science Education*, *90*(4), 681-706.

Vesterinen, V-M., Manassero-Mas, M-A., & Vázquez-Alonso, Á. (2014). History, Philosophy, and Sociology of Science and Science- Technology-Society Traditions in Science Education: Continuities and Discontinuities. In M. R. Matthews (ed.), *International*

Handbook of Research in History, Philosophy and Science Teaching (pp. 1895-1925). Dordrecht: Springer.

Wan, Z. H., Wong, S. L. & Zhan, Y. (2013). Teaching nature of science to preservice science teachers: A phenomenographic study of Chinese teacher educators' conceptions. *Science & Education*, *22*, 2593–2619.

Wong, S. L., Hodson, D., Kwan, J. & Yung, B. H. W. (2008). Turning Crisis into Opportunity: Enhancing student-teachers' understanding of nature of science and scientific inquiry through a case study of the scientific research in severe acute respiratory syndrome. *International Journal of Science Education*, *30*, 1417-1439.