Social Construction of Technology: An Experience for Development of Critical-thinking and Nature of Science and Technology

Vanessa Ortega-Quevedo[a], Cristina Gil-Puente[a], and Cristina Vallés Rapp[a]

Corresponding author:
[a] Department of Sciences Education, Social Sciences Education and Mathematics Education, Universidad de Valladolid, Plaza Universidad 1a, 47002, Valladolid, Spain

Keywords: critical thinking, science and technology, elementary education/primary education

Abstract

This research aims to contribute to the development of critical thinking skills and concepts of the nature of science and technology through joint work based on primary school curriculum content (energy). It has a mixed design in methodological terms, applying a variety of techniques such as surveys, interviews, participant observation and documentary analysis. The total sample of participants is 130 students aged 11-12 from five different schools. The results of the study show improvements in the concepts of the nature of science and technology (dependence on the use of new technologies and control of technological development by individuals), as well as in critical thinking skills (thinking as hypothesis testing and argument analysis), among participants. For example, there is a perceived increase of 0.34 (2-point scale) in the mean between the initial and final assessment of dependence on the use of new technologies This leads us to conclude that the teaching design implemented is effective for improvement in both areas.

Introduction

Contemporary society feeds on scientific and technological developments, often to the point of dependence. However, the scientific and technological literacy of the general public has often been considered insufficient throughout the decades (Pleasants, Clough, Olson, & Miller, 2019). Promoting scientific and technological literacy at the global level is therefore essential to understand and critically evaluate the role of science and technology in everyday life (Acevedo, García-Carmona, & Aragón, 2017; Vázquez & Manassero, 2012).

Moreover, the challenges faced by our planet make the need for this kind of literacy more urgent. The current state of environmental alert or situations like the COVID-19 pandemic highlight the need for training in science and technology if we are to understand the scope and relevance of decisions based on scientific and technological knowledge that affect society. Two elements are key to scientific and technological literacy: Understanding the Nature of Science and Technology (NoSaT) and acquiring Critical Thinking (CT) skills.

In response to social demands, this study seeks to contribute to the development of CT skills and of concepts of NoSaT through joint work based on curriculum content. We understand NoSaT as the integration of nature of science (NoS) and nature of technology (NoT) into a single construct, while keeping their elements separate to observe the interactions between science and technology, and between them and society (Vázquez & Manassero, 2012). A teaching-learning sequence (TLS) is proposed, designed to develop CT skills while facilitating the learning of NoSaT concepts by focusing on high-impact curriculum content such as energy (definition, properties, manifestations, renewable and non-renewable sources, environmental impact and fracking). On the other hand, CT is considered to be one of the competences for the twenty-first century. PC skills are a set of knowledge that can be transferred to the real world in order to function in contemporary society (Almerich et al., 2020, Mercado-Ramirez, 2021). They are a set of skills that enable people to develop strategies and habits to implement effective thinking processes (making good decisions, building evidence-based arguments, analysing information, and so on) (Ritchhart, Church, & Morrison, 2014).

In particular, the specific objectives are:

- To test whether conceptions of NoSaT related to the social construction of technology improve after the implementation of a TLS. Specifically, about:
    - Dependence on the use of new technologies
    - Control of technological development by individuals
- Check for improvement of PC capabilities (argument analysis and thinking as hypothesis testing) after the intervention.

Since both NoSaT and CT are broad constructs, the teaching design addresses specific concepts of NoSaT and CT skills, for this is the only way of making them explicit. Definitions of CT and NoSaT are therefore in order here, before defining the elements chosen for analysis.

Concepts of Nature of Science and Technology

In this study, NoSaT is used as an umbrella term in education, encompassing both NoS and NoT (Vázquez, Manassero, & Talavera , 2010). It is defined as metaknowledge about science and technology, arising from interdisciplinary reflections by philosophers, historians, sociologists of science and technology, scientists and science educators. From this perspective, the term includes epistemic and non-epistemic aspects, as well as science, technology and society (STS) factors (Acevedo, Aragón, & García-Carmona, 2018).

Manassero and Vázquez (2019) present a taxonomy that organises the different elements that make up the NoSaT construct. This taxonomy proposes four basic aspects as organisers: definitions and interactions between science and technology; external sociology of science and technology; internal sociology of science and technology; and epistemology. In this study we focus on components of the organiser called "internal sociology of science and technology". In this way we can say that the internal sociology of science in turn is made up of elements related to characteristics of scientists; social construction of scientific knowledge; and social construction of technology (technological decisions; autonomy of technology). In particular, due to the age of the students and the time allocated to the intervention, we will focus on learning elements related to the social construction of technology, such as dependence on the use of new technologies; and control of technological development by individuals.

Finally, it should be noted that although from this study we understand a holistic view of NoSaT, but for its work in specific classroom situations at an early age, we advocate the teaching of specific elements that are part of each of the dimensions of NoSaT. In this way, throughout schooling, knowledge will be acquired which, at higher levels, will serve as a foundation for understanding the construct.

Critical Thinking skills

The CT construct is defined and bounded differently depending on the discipline and the author (Ennis, 1996; Halpern 2014; Paul, 2005). What the various definitions have in common is that

CT is a set of intentional processes triggered to draw conclusions and establish how these conclusions were reached. When using CT, related data or problems to be solved are broken down, summarised and thoughtfully evaluated in order to come to a conclusion or find a solution. The conclusion or solution is in turn analysed to see if it can be improved. (Ortega-Quevedo, Gil-Puente, Vallés, & López-Luengo, 2020, p. 94)
Authors tend to agree that CT has five main components: abilities, dispositions, attitudes and values, rules, and knowledge (Tenreiro-Vieira and Vieira, 2021). According to Halpern (1998), the cognitive component of CT has five dimensions: argument analysis, thinking as hypothesis testing, verbal reasoning, using likelihood and uncertainty, and decision making and problem solving. This study focuses on argument analysis and thinking as hypothesis testing. However, it should be borne in mind that all five dimensions are closely related and cannot be dealt with separately. This means that even though only two of them have been selected, the other three will also be taken into consideration.

Argument analysis: The ability to think by breaking down arguments into their components and distinguishing between reasons, assumptions, qualifiers, counterarguments and conclusions (Halpern, 2014).

Thinking as hypothesis testing: A hypothesis is a supposition or an explanation about the world around us made on the basis of limited evidence, which needs to be proved or disproved. By testing a hypothesis, we get to know the truth about something. It is a way of building knowledge in order to interact with our environment (Halpern, 2014).

Study design

Since a variety of research techniques were required in this study, a longitudinal multi-method design was used (three months' time), comprising both quantitative and qualitative research techniques to enable data triangulation and to offset the shortcomings of one method with the benefits of another. Within this framework, the research was conducted by collecting the data through teacher/researcher interventions in more than one classroom (Figure 1).

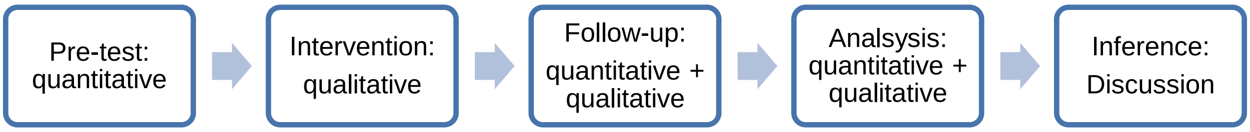

Figure 1. Phases of the multi-method design

Data collection tools and techniques

Techniques such as survey, interview and documentary analysis were used and data were collected through instruments such as the opinion survey on science, technology and society (OSSTS) Adaptation and the Critical Thinking Assessment instrument, the teacher/researcher's diary and students' productions.

Adaptation of the OSSTS validated by Ortega-Quevedo and Gil-Puente (2019).
This instrument, designed to be implemented with students aged 11 to 12 years, is applied in the first and third phase of the research as a pre- and post-test of the participating students. The application was carried out in the classroom of each of the class-groups. Specifically, the results obtained in items 80131 and 80211 are analysed (Figure 2).

Item 80131 adaptation:

When a new technology emerges (e.g., a new computer or a new drug against cancer), it can then be implemented or not.

Some participant answers to be considered:

- The decision to apply new technologies depends mostly on its benefits to society. If it has too many drawbacks, it will be rejected by society and this can stop further development.
- The decision depends not only on the advantages and disadvantages of the new technology, but also on how well it works, how much it costs or how effective it is.

Item 80211 adaptation:

Technological development can be controlled by individuals.

Some participant answers to be considered:

- Yes, because scientists and technologists are individuals too. If they control technological development, it is also individuals who are in control.
- No, because technological advancements are so fast-paced those individuals cannot keep abreast of them.

Figure 2. OSSTS adaptation items

Critical Thinking Assessment Instrument validated by Ortega-Quevedo and Gil-Puente (2019)

This instrument is implemented with the students and is applied in the first and third research phase (pre- and post-test) in the classroom of each of the groups-classes. Specifically, the results obtained in the skills of argumentative analysis and verbal reasoning are analysed (Figure 3).

Scenario for the assessment of argument analysis:

There are many opportunities for computer experts. It is a good area to specialise in, offering interesting jobs, good career prospects and high salaries. Of course, if you do not like maths or are not much of an indoor person, it would not be the right choice for you.

Students are asked to identify reasons, counterarguments, and conclusions in this statement.

Scenario for the assessment of thinking as hypothesis testing:

Following a televised debate on the use of video games, viewers were encouraged to go to the TV channel website and take a survey: Were they for or against children playing video games? About 1000 people took the survey in the first hour, with the same number of votes for and against the idea. On the following day, the host announced the survey results, concluding that the people in the country were equally divided in their opinions about the use of video games by boys and girls.

Students are asked to say if statements like the ones below are true or false:
- The speakers for and against the use of video games in the debate were equally convincing.
- The viewers who took the survey may not be representative of the people in the country.

Figure 3. Situations of the critical thinking assessment tool

Classroom diary
The class diary used is kept by the teacher/researcher during the implementation phase and is of an analytical type, as the aim is to analyse specific aspects within a given observation context (Zabalza, 2004).

Rubrics
A rubric or descriptive scale is a scale that establishes different levels of achievement and, in each of these levels, a description is developed that is as precise as possible about the characteristics of the exercise to be evaluated (López-Pastor & Pérez-Pueyo, 2017). This instrument is used in the assessment of children's productions (available in the Supplementary Material).

Methods of analysis
The quantitative analysis of the data was performed using the statistical software tool SPSS 23. The Kolmogorov-Smirnov test was used for testing normality, showing that some of the variables were not normally distributed (p-value less than .05). The Wilcoxon signed-ranked test was used to compare pre-test and post-test results, while the Mann-Whitney U test was used to compare significance in improvements for each group vis-à-vis the control group.

As for qualitative analysis, a system of categories was implemented to analyse classroom journals and interview transcripts (see Table 1). This categorisation has been constructed

deductively, taking as a reference Manassero and Vázquez (2019) in the category of NoSaT and Halpern (2014) in the category of CP. The analysis of the children's productions (documentary analysis) was carried out using evaluation rubrics. These rubrics were developed based on theoretical references such as Vázquez and Manassero (2012) and Halpern (2014).
In order to ensure the confidentiality of participating subjects and schools, codes were assigned to the data collection documents according to participant number, gender, school number, group and level of learning (high, average, 'in progress').

Table 1. Categories for qualitative analysis

Categories		Subcategories
NoSaT improvements	Social construction of technology	Dependence on the use of new technologies Control of technological development by individuals
CT improvements	Argument analysis	Use of reasons and counterarguments Formulation of conclusions
	Thinking as hypothesis testing	Formulation of evidence-based hypotheses Testing and correction of hypotheses

Sampling
The participants in this study included 130 sixth-grade elementary school students (aged 11-12), who were part of 7 groups in 5 schools in the province of Segovia, Spain (Table 2). One of the 7 groups was the control group. These schools and groups were chosen according to the so-called convenience sampling. In order to gain access to the schools, the researcher requested permission from the educational inspection (evaluation by a committee of the relevance of the study, the didactic and assessment instruments implemented and the impact of the proposal on the students' education; approval 2016/679).

The validity of the control group was determined using the Kruskal-Wallis H test, showing p-values of less than .05 for all variables and thus proving the statistical equivalence between the control group and the experimental groups in terms of initial knowledge. Finally, each group's openness to the teaching methodology was assessed in informal interviews with the regular teachers.

Table 2. Participants in the study by schools and groups

Groups	Participants			Openness to methodology
	Boys	Girls	Total	
School_1_Group_A	12	10	22	Average
School_1_Group_B	12	9	21	Average
School_2_Group_A	9	10	19	Average
School_3_Group_B	11	8	19	High
School_4_Group_A	7	11	18	Low
School_4_Group_B	5	13	18	Low
School_5_Control_Group	9	4	13	
	65	65	130	

Note 1: Affinity with the methodology is established through informal conversations with natural classroom teachers, through which it is clarified which teaching models and methodologies the teacher usually uses.
Note 2: The socio-economic profile of the participating families is established through the analysis of the School Education Projects and informal interviews with the natural teachers of each class-group.

Finally, it is worth highlighting the sampling carried out to conduct the interviews with the students. In this case, it is determined that three interviews will be conducted within each group-class on an experimental basis, as this is the appropriate number for research that seeks to test hypotheses between different groups (Kvale, 2011). On this basis, the natural teachers in each group-class were asked to select one student per learning level (high, average, ‘in progress’) to participate in these interviews.

Teaching intervention

A special TLS was designed for 11- and 12-year-old students, with activities dealing with energy-related content (readers can find examples of activities in the Supplementary Material) and also with the CT skills and NoSaT concepts selected (Table 3).

Table 3. TLS content by sessions

Session	NoSaT content	CT content	Curriculum content
1	Dependence on the use of new technologies	Argument analysis	Concept, properties and types of energy
2	Dependence on the use of new technologies and control of technological development by individuals	Argument analysis and thinking as hypothesis testing	Sources of renewable and non-renewable energy
3	Dependence on the use of new technologies and control of technological development by individuals	Argument analysis and thinking as hypothesis testing	Fracking

Note: more information on TLS can be found at Ortega-Quevedo et al. (2020).

In designing the TLS, it was considered that an explicit approach was most effective in improving NoSaT concepts and contributing to the development of CT skills (Acevedo, 2009; Tenreiro-Vieira & Vieira, 2021). Every TLS session consisted of three activities:

- Background knowledge activation through thinking routines (Ritchhart, Church, & Morrison, 2014), followed by group sharing of ideas.
- Listening-questioning dialogue (presentation of content).
- Knowledge transfer (controversial issues to discuss in pairs first and then by the group).

Learning assessment was also considered as key to the TLS (López-Pastor & Pérez-Puello, 2017), since it is the element that brings innovation (development of TC skills and NoSaT concepts), learning and the transfer of curriculum content (energy) together. According to Dochy, Segers, and Dierick (2002), without the corresponding adapted assessment, innovations like the one in this study are meaningless as educational processes.

Results and discussion

NoSaT improvements: Social construction of technology and dependence on the use of new technologies

In the first place, the results of item 80131 in the adaptation of OSSTS were analysed. Figure 3 shows the graphic representation of the different variables of the study, highlighting the contrasting scores of the control group and the experimental groups at the various measurement phases, as well as the contrast between the control group and the groups exposed to the teaching intervention. The score went from -1 to 1, 1 being the top score.

As shown in Figure 4, there is significant change in the distribution of results between the pre-test and post-test phases for the experimental groups, both individually and collectively. On the contrary, the score for the control group is lower in the post-test measurement.

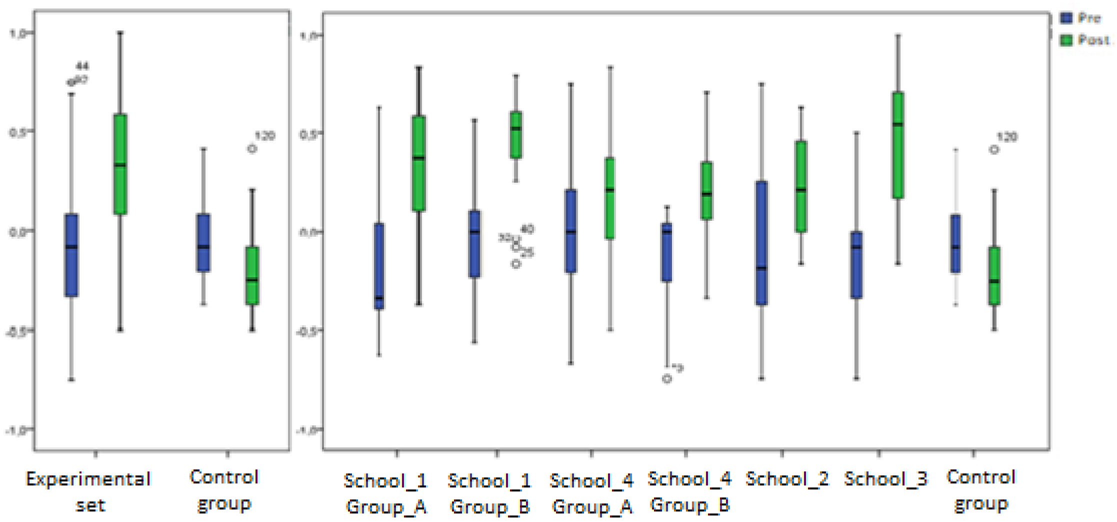

Figure 4. OSSTS item 80131 adaptation graphics

The Wilcoxon signed-ranked test was used to test the significance of the differences in pre-test and post-test results for each group. It was concluded that, with a p-value of .000 between the medians in the different phases for the experimental groups, the difference was significant, with an error of .05. Broken down, all the experimental groups showed p-values of < .05 except School 4 Group A, whose p-value was .118. Again, these values mean that differences are significant for all groups except School 4 Group A. As for the control group, the p-value was .07, which means that the medians are the same.

Secondly, the Mann-Whitney U test was used to see if the improvements in each group were significant as compared to the control group. The comparison of the medians for the experimental groups taken together and the control group showed a p-value of .000, which means that the improvements in the experimental groups are significant, with a margin of error of 5 percent. When analysing the improvements for each group (using the same instrument), the p-values were < .05, which means that the difference between the medians is significant.

The analysis of classroom journals showed that reflection on the use of new technologies began during the consolidation activities in the three sessions, when the students abandoned the idea that all technologies are developed and implemented only for the benefit of society. Moving away from this idea, they began to consider other factors, like economic interests, as well as both the advantages and disadvantages of implementing new technologies. The ensuing discussion, as recorded in the journals, shows how the students began to think about issues that they had not thought about before, such as the interests of the private companies promoting the implementation of power stations or the exploitation of energy resources, or the impact of the implementation or exploitation in question. Even though the discussion was not too deep, given the students' developmental stage, it came close to some of the ideas gathered as adequate or general explanations for some of the questions posed by Torres and Solbes (2016).

Some of these ideas were present in the activities carried out. When answering questions like 'What does the implementation of new technologies depend on?', the students gave answers that came close to the standard (Vázquez & Manassero, 2012) – *'Venefits*[sic], *money, companies, people'* –, mostly plausible answers – *'Taking both adbantages*[sic] *and disadbantages*[sic] *into account'* – and a few wrong answers – *'On the people who make them or on scientists.'* In the evaluation of the student productions using the rubrics designed for the study, out of 52 exercises reviewed (documents written in small groups), 6 had a low score, 30 achieved the average and 16 got a high score. Moreover, 46 of the 52 groups produced acceptable documents, which is in line with the improvements observed in the quantitative analysis performed and the classroom discussions analysed.

Finally, during the follow-up one month after the intervention, the students were asked whether they had changed their minds about what the implementation of new technologies depends on. The students with a high level of learning adequately explained that they still thought the same because they were already familiar with the topic discussed in class. On the other hand, some of the students with a level of learning in progress tried to share their opinions, even if they were inadequate. For instance:

> 83.M.5.A.A.: *No, it depends on the benefits to society, on the time it takes, on the jobs it creates…*
> 46.H.3.B.P.: *Yes.*

NoSaT improvements: Social construction of technology and control of technological development by individuals

Figure 5 shows the main descriptive statistics for the results of item 80211 in the adaptation of the OSSTS. A higher score can be observed in the post-test phase for the experimental groups, both together and individually. On the contrary, the control group shows no difference between scores.

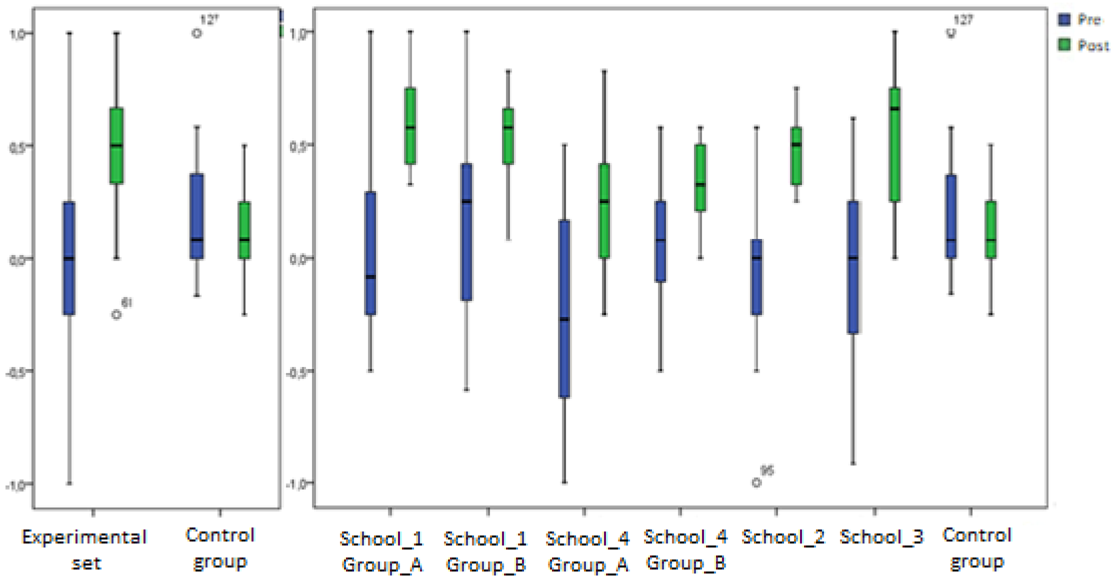

Figure 5. OSSTS item 80211 adaptation graphics

The Wilcoxon signed-ranked test was used to test the significance of the difference in pre-test and post-test results. With a p-value of .000 for the experimental groups as a whole, it can be said with 95 percent confidence that the difference between the medians in the first and second measurements is significant. Moreover, when broken down, the experimental groups showed p-values of less than .05, which confirms the previous finding. In the case of the control group, the p-value was .39 thus no improvement is observed.

Then, the Mann-Whitney U test was used to see if the improvements in each group were significant as compared to the control group. The p-value for the difference in the medians was .000, which means the difference is significant. When comparing class groups with the control group, the p-values were < .05 for all groups except School 4 Group B. This confirms the improvements for all individual groups but School 4 Group B.

The classroom journals revealed that the students initially believed that the development of new technologies was controlled by scientists and engineers alone. However, after the discussions that were part of the consolidation activities in the last two sessions, they reached conclusions like: *'It's the companies with the money that control technology'* (classroom transcript). Likewise, they expressed their disagreement with this state of affairs, adding that they would like to have a say in it.

The thoughts jotted down in the journals were consistent with the results of the activities. The students shared their opinions on the issue of control of technological development by individuals in the form of adequate answers to the question – *'No. It's controlled by the government and businesses'* –, plausible ideas – *'Many people should sign a petition to stop this'* – and mistaken concepts – *'Yes, because we use the technology.'* In the evaluation of the student productions using the rubrics designed for the study, out of 52 exercises reviewed, 5 had a low score, 29 achieved the average and 18 got a high score.

Finally, in the follow-up interviews, the students were asked whether they had changed their minds about the control of technological development by individuals. Most of them said they had, their opinions now being close to what experts consider to be right (Vázquez & Manassero, 2012). However, there was at least one in each level of learning who continued to hold an erroneous view or was unable to express their ideas adequately. For example:

> 83.M.2.A.A.: *Yes, I used to think that scientists controlled technology because they developed it, but now I think it's those who have the money. There were things in the survey that made me think it over.*
> 25.M.1.B.P.: *No. I don't know, really.*

CT improvements: Argument analysis, use of reasons and counterarguments, and formulation of conclusions

Figure 6 shows the main descriptive statistics. A higher score can be observed in the post-test phase for the experimental groups, both together and individually. On the contrary, the control group shows a lower post-test score.

The Wilcoxon signed-ranked test was used to test the significance of these observations. With a p-value of .000 for the experimental groups as a whole, it can be said that the difference between the medians in the first and second measurements is significant. When broken down, all experimental groups except School 2 showed p-values of < .05. School 2 had a high initial level, so there was little difference in pre-test and post-test scores. As to the control group, the p-value was .56 and thus no improvement is observed.

Then, the Mann-Whitney U test was used to see if the improvements in each group were significant as compared to the control group. When comparing the experimental groups as a whole with the control group, the p-value was .02, which means that the difference between the medians is 95 percent significant. In the comparison between class groups and the control group, the p-value was .05 for 3 groups (School 1 Group A, School 4 Group B and School 3). Based on these results, the difference between the medians was established for 3 of the 6 experimental groups.

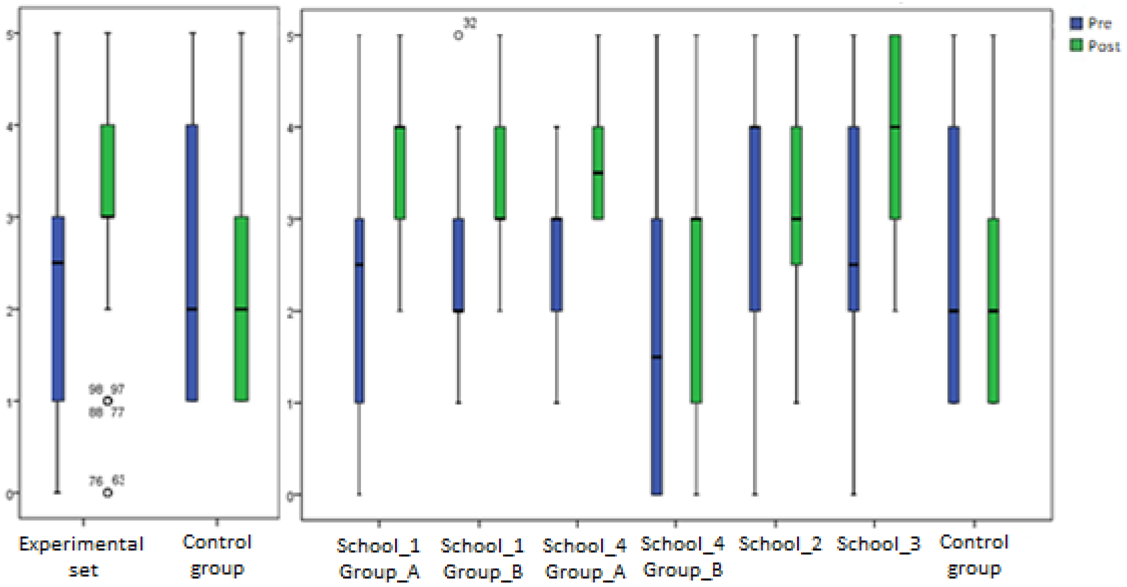

Figure 6. Argument analysis assessment graphics

To sum up, the quantitative results show significant improvements between the various measurement phases and, in some cases, vis-à-vis the control group. This positive development is consistent with the findings in the study by Porras, Tuay, and Ladino (2020), including an intervention with similar goals and evaluation procedures with a group of 15 to 17 year old students.

The classroom journals show how the groups began to work on argumentation in the first session, and how both argument-related terminology and argument building processes were present throughout the TLS. They also show how constant practice enabled the students to build increasingly complex arguments, featuring better-defined reasons and counterarguments, and to reach more sophisticated conclusions.

These findings are consistent with the results of the activities, where the rubric-based evaluation of the exercises shows that the arguments improved as the TLS unfolded (Figure 7). The analysis of this development indicates a lower number of student assignments with low-level achievement and a higher number of high-achieving assignments. These results coincide with improvements obtained in other studies that seek to improve students' argumentative processes through energy-related topics, such as the study of Skoumios and Balia (2021)

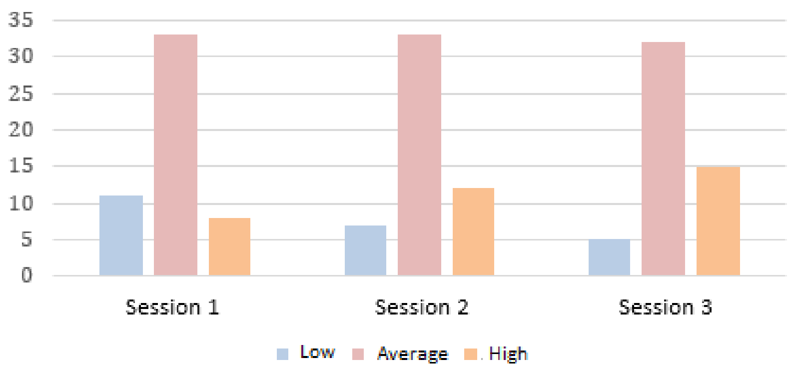

Figure 7. Argument analysis production results

Finally, in the follow-up interviews, the students were asked if they remembered the video introduced in the third session. Most of them answered they did. This means that they were able to identify and remember the elements of the argument, which is consistent with the observations about improvements. A few examples:

> 104.H.3.B.A.: *Yes, a group agreed, saying that no harm was done, that it was under control and that it would boost the economy and energy development, while others were against the idea because of the negative impact and river pollution.*
>
> 102.M.3.B.M.: *Yes, a group said the effects on the environment would be disastrous, that the whole thing would be too dangerous, but others said quite the opposite.*

CT improvements: Thinking as hypothesis testing, formulation of evidence-based hypothesis, and testing and correction of hypotheses

Figure 8 shows the main statistics for the assessment of thinking as hypothesis testing. The scores observed were low (top score 5), even though they were higher in the post-test stage. A different performance was observed in the control group.

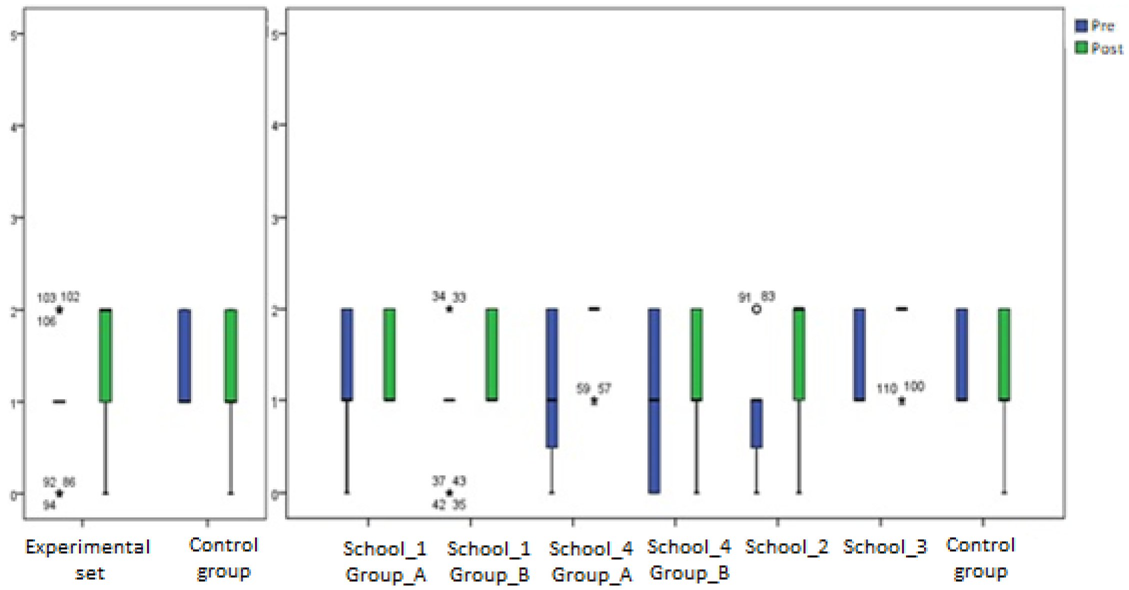

Figure 8. Thinking as hypothesis testing assessment graphics

The Wilcoxon signed-ranked test was used to test the difference in measurement results. When comparing the results for the experimental groups taken together, the p-value was .000. Thus, it can be said with 95 percent confidence that the students improved their skills for thinking as hypothesis testing. The comparison for individual groups shows p-values of < .05 in all groups except School 1 Group A and School 4 Group B, which means the improvements are significant in 4 of the 6 experimental groups. The p-value for the control group was .16, so the medians can be said to be equal.

Then, the Mann-Whitney U test was used to see whether the improvements in each group were significant as compared to the control group. When comparing the experimental groups as a whole with the control group, the p-value was .07, which means that the difference between the medians is not significant. In the comparison between the control group and individual experimental groups, the p-value was .05 for 2 groups (School 4 Group A and School 3). Based on this, the improvements in student results as compared to the control group were established as significant for 2 schools.

The analysis of classroom journal content showed that the concept of hypothesis was introduced in the first session. After this, the students worked with hypothesis testing in all the activities, considering whether the hypotheses were applicable to the population at large or to a sample only. By regularly applying these thinking processes, they gradually became aware of them as strategies and began to identify when they were being used. In the last session, the students were able to apply these processes, analysing rationale soundness, credibility or the number of data sources used, when testing their initial hypotheses.

This could be seen in the students' workbooks, when they introduced changes in the exercises, rephrasing or correcting the initial hypotheses following classroom discussions. On the basis of the rubric-based evaluation of the workbooks, a bar chart was used to show the results of the three sessions (Figure 9). This chart shows that the number of groups with low levels of achievement decreases as the sessions unfold (from 23 in the first session to 12 in the third).

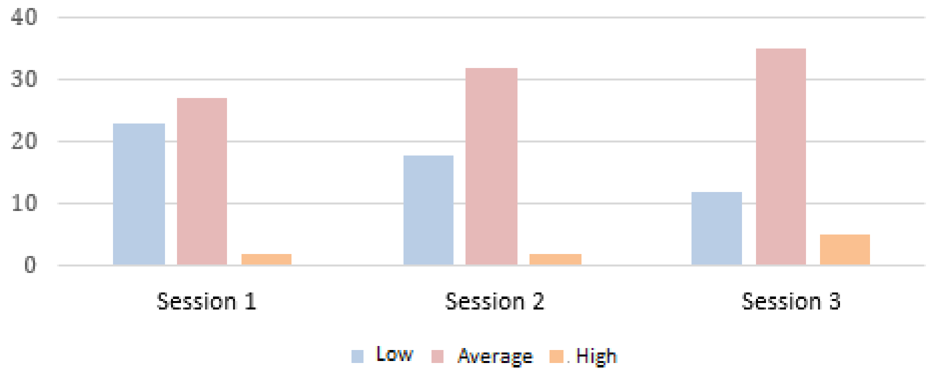

Figure 9. Results of student productions on thinking as hypothesis testing

Finally, in the follow-up interviews, the students were asked if, based on the height of the children in their school, they could say that all 11-year-olds in Segovia were tall. The students with a high or average level of learning proffered explanations about generalisations from a sample to the entire population, showing that they were applying valid generalisation and information checking processes. On the other hand, the students with a level of learning in progress gave only few reasons why the hypothesis could not be tested. A few examples are shown below.

> 83.M.2.A.A.: *No, because we don't represent all the children in Segovia. If you went to every school and do the maths, maybe you could say that.*
>
> 102.M.3.B.M.: *No, because we're not all equally tall.*

Conclusion

The results of the study enable us to conclude that the teaching proposal designed and implemented for this research contributes to the improvement of critical thinking (CT) skills (argument analysis and thinking as hypothesis testing) and to the development of concepts of the nature of science and technology (NoSaT- Social construction of technology: dependence on the use of new technologies and control of technological development by individuals) in 11- and 12-year-old schoolchildren. The quantitative analysis using the Wilcoxon signed-ranked test showed that there were significant improvements when comparing pre-test and post-test measurements for the experimental groups taken together. The qualitative data support this finding, introducing nuances in the ways in which the students answered NoSaT questions, corrected their initial hypotheses (student productions) and learned to build more complex arguments.

When broken down by group, the analysis shows greater improvement in School 4, although in some cases this is not significant when comparing between phases or with the control group. This can be related to the groups' low openness to the methodology, as they were used to more traditional teaching-learning processes and were unfamiliar with dialogue in the classroom. Thus, the proposal has potential to promote CT skills and improve NoSaT concepts among students at this stage of education.

The study addresses elements of impact for research in science didactics in a context that is little contemplated, such as primary education. Consequently, the study provides data on how the improvement of different thinking skills and content on the internal sociology of science can begin to be worked on at this stage. While it is true that the nature of science has been trying for years to be integrated into educational classrooms, including primary education, this integration is still far from being achieved. However, publishing LSAs that demonstrate the learning of NoSat issues allows the principles on which they are based to be made available to the scientific and educational community and can thus be integrated into syllabuses, textbooks, teacher training courses etc.

Finally, this study has some limitations that should be highlighted. Firstly, students were not randomly selected, and the sample used could not be representative of the entire reality of Spain, but it is quite representative of the population of Segovia. Although, in general, it is difficult to access school classrooms and students in Spain, it would be interesting expanding the sample to other regions and contexts. In contrast, the strength of the study is the extensive methodological triangulation that contrasts the results obtained from four different types of data collection instruments. In addition, data are provided in a less studied context, primary education, and all interventions have been carried out by the same teacher, eliminating extraneous variables related to teaching style.

References

Acevedo, J.A. (2009). Enfoques explícitos versus implícitos en la enseñanza de la naturaleza de la ciencia. *Revista Eureka sobre Enseñanza y Divulgación de las Ciencias, 6*(3), 355-386.

Acevedo, J.A., Aragón, M., & García-Carmona, A. (2018). Comprensión de futuros profesores de ciencia sobre aspectos epistémicos de la naturaleza de la ciencia en cuatro controversias de historia de la ciencia. *Revista Científica, 33*(3), 344-355. https://doi.org/10.14483/23448350.13355

Acevedo, J.A., García-Carmona, A., & Aragón, M. (2017). *Enseñar y aprender sobre naturaleza de la ciencia mediante el análisis de controversias de historia de la ciencia.* Organización de Estados Iberoamericanos para la Educación, la Ciencia y la Cultura

Almerich, G., Suárez-Rodríguez, J., Díaz-García, I., & Orellana, N. (2020). Estructura de las competencias del siglo XXI en alumnado del ámbito educativo. Factores personales influyentes. *Educación XX1, 23*(1), 45-74. https://dx.doi.org/10.5944/educXX1.23853

Dochy, F., Segers, M., & Dierick, S. (2002). Nuevas vías de aprendizaje y enseñanza y sus consecuencias: una nueva era de evaluación. *Docencia universitaria, 2*(2), 1-29.

Ennis, R.H. (1996). *Critical Thinking*. Prentice-Hall.

Halpern, D.F. (1998). Teaching critical thinking for transfer across domains. *American Psychologist, 53*(4), 449-455. https://doi.org/10.1037/0003-066X.53.4.449

Halpern, D.F. (2014). *Thought and knowledge. An Introduction to Critical Thinking*. Psychology Press.

Kvale, S. (2011). *Las entrevistas en Investigación Cualitativa*. Morata, S. L.

López-Pastor V. M., & Pérez-Pueyo, A. (coords.) (2017). *Buenas prácticas docentes. Evaluación formativa y compartida en educación: experiencias de éxito en todas las etapas educativas*. Universidad de León

Manassero, M.A., & Vázquez, A. (2019). Conceptualización y taxonomía para estructurar los conocimientos acerca de la ciencia. *Revista Eureka Sobre Enseñanza Y Divulgación De Las Ciencias, 16*(3), 3104-3117. https://doi.org/10.25267/Rev_Eureka_ensen_divulg_cienc.2019.v16.i3.3104

Mercado-Ramirez, H. J. (2021). Facilitating Computer-Supported Collaborative Learning with Question-Asking Scripting Activity and its Effects on Students' Conceptual Understanding and Critical Thinking in Science. *International Journal of Innovation in Science and Mathematics Education, 29*(1), 31-45. https://doi.org/10.30722/IJISME.29.01.003

Ortega-Quevedo, V., & Gil-Puente, C. (2019). La naturaleza de la ciencia y la tecnología. Una experiencia para desarrollar el pensamiento crítico. *Revista Científica, 35*(2), 167-182. https://doi.org/10.14483/23448350.14095

Ortega-Quevedo, V., Gil-Puente, C., Vallés, C., & López-Luengo, M.A. (2020). Diseño y validación de instrumentos de evaluación de Pensamiento Crítico en Educación. *Tecné, Episteme y Didaxis, 48*, 91-110. https://doi.org/10.17227/ted.num48-12383

Paul, R. (2005). The state of critical thinking today. *New Directions for Community Colleges*, 27-38.

Pleasants, J., Clough, M.P., Olson, J. K., & Miller. G. (2019). Fundamental Issues Regarding the Nature of Technology. *Science & Education*, *28*, 561–597. https://doi.org/10.1007/s11191-019-00056-y

Porras, Y., Tuay, N., & Ladino, Y. (2020). Desarrollo de la habilidad argumentativa en estudiantes de educación media desde el enfoque de la Naturaleza de la Ciencia y la Tecnología. *Tecné, Episteme y Didaxis, 48*, p143-161.

Ritchhart, R., Church, M., & Morrison, K. (2014). *Hacer visible el pensamiento. Cómo promover el compromiso, la comprensión y la autonomía de los estudiantes*. Paidós.

Skoumios, M., & Balia, C. (2021). The Impact of Teaching Interventions for Electrical Circuits on the Structure of Primary School Students' Written Arguments. *International Journal of Innovation in Science and Mathematics Education*, *29*(1), 16-30 https://doi.org/10.30722/IJISME.29.01.002

Tenreiro-Vieira, C., & Vieira, R.M. (2021). Promover o pensamento crítico e criativo no ensino das ciências: propostas didáticas e seus contributos em alunos portugueses. *Investigações em Ensino de Ciências*, *26*(1), 70-84, https://doi.org/10.22600/1518-8795.ienci2021v26n1p70

Torres, N.Y., & Solbes, J. (2016). Contribuciones de una intervención didáctica usando cuestiones sociocientíficas para desarrollar el pensamiento crítico. *Enseñanza de las Ciencias, 34*(2), p43-65. https://doi.org/10.5565/rev/ensciencias.1638

Vázquez-Alonso, A., & Manassero, M.A. (2012). La selección de contenidos para enseñar naturaleza de la ciencia y tecnología (parte 1): Una revisión de las aportaciones de la investigación didáctica. *Revista Eureka sobre Enseñanza y Divulgación de las Ciencias, 9*(1), p2-31. http://doi.org/10.25267/Rev_Eureka_ensen_divulg_cienc.2012.v9.i1.03

Vázquez, A., Manassero, M.A., & Talavera, M. (2010). Actitudes y creencias sobre naturaleza de la ciencia y la tecnología en una muestra representativa de jóvenes estudiantes. *Revista Electrónica de Enseñanza de las Ciencias*, *9*(2), p333-352. http://reec.uvigo.es/volumenes/volumen9/ART3_Vol9_N2.pdf

Zabalza, M.A. (2004). *Diarios de clase. Un instrumento de investigación y desarrollo profesional*. Narcea.