Cite this article:

García-Carmona, A. (2019). Pre-service primary science teachers' abilities for solving a measurement problem through inquiry. *International Journal of Science and Mathematics Education*, *17*(1), 1-21. https://doi.org/10.1007/s10763-017-9858-7

Pre-service primary science teachers' abilities for solving a measurement problem through inquiry

Antonio García-Carmona

Departamento de Didáctica de las Ciencias Experimentales y Sociales, Universidad de Sevilla, Spain.

Email: garcia-carmona@us.es; ORCID: https://orcid.org/0000-0001-5952-0340

Abstract. This article presents a qualitative descriptive-interpretive study about the abilities of pre-service primary science teachers (PPTs) to solve a measurement problem put to them as an activity of scientific inquiry. The participants in the study were 23 PPTs receiving training in science teaching, organized into small groups to solve the proposed problem. The data were collected through reports they prepared from a script with open-response questions that referred on the one hand to the planning, development, and conclusions obtained with the scientific inquiry, and on the other to promoting the PPTs' metacognitive reflection regarding the difficulties encountered, the lessons learned, and their proposals for improvement for future activities on similar measurements. The information was analysed in three phases to increasingly refine the coding of this. The results revealed that, overall, the PPTs were able to complete the scientific inquiry with a fairly acceptable degree of efficacy. A discussion of the results emphasizes the usefulness of the procedural-type scientific inquiry activities to initiate PPTs in their approach to inquiry-based science education.

Keywords: experimental activity; initial teacher training; inquiry-based science education; measurement; primary education

Introduction

For years, inquiry-based learning has been considered to be the most appropriate educational approach to learning science (Abd-El-Khalick et al., 2004; Rocard et al., 2007). Within this approach, especially interesting are the experimental activities (ExA's) with which students can observe, investigate, and develop an understanding of the physical world surrounding them through direct interaction with phenomena and the manipulation of objects and materials (Criado & García-Carmona, 2011). Thus, the students are encouraged to develop basic scientific abilities to learn science by doing science (Hodson, 2014).

However, despite the fact that the literature highlights the educational benefits of ExA's for learning science through inquiry (Miner et al., 2010), the reality is that their inclusion in science classes is still scarce and usually through superficial approaches far removed from inquiry-based learning (Flores et al., 2009). This has recently been confirmed in the actual science education promoted in Spanish primary education classrooms (Cañal et al., 2013). One of the main reasons for this situation is that pre-service primary science teachers (PPTs) do not usually receive any adequate training in inquiry-based educational strategies (Cañal et al., 2011), nor do they have opportunities to test such strategies during their practicum periods in schools (Cortés et al., 2012; García-Carmona & Cruz-Guzmán, 2016).

In view of this, for some years in our field we have been developing a PPT training project about the design and implementation of ExA's as inquiry (Criado & García-Carmona, 2011). We have already reported an analysis of the PPTs' pedagogical beliefs on the use of ExA's to learn science (García-Carmona et al., 2016), the difficulties that they have in planning the complete development of a scientific inquiry (García-Carmona et al., 2017), and the types of inquiry questions they propose when designing ExA's (Cruz-Guzmán et al., 2016).

Since the best way for PPTs to assimilate inquiry-based science learning strategies is for them to experiment with them first as though they themselves were science students (Newman et al., 2004), at present we are focusing on the PPTs acquiring abilities to conduct a complete scientific inquiry in the context of an ExA. The purpose of this paper is to present the results and conclusions of a study about the abilities the PPTs show when they have to solve a measurement problem through scientific inquiry.

Review of the literature

Experimental activities in science education

The literature shows that the ExA's normally posed in science classes consist of simple observations of phenomena with the manipulation of some instruments and materials, following step-by-step instructions as if it were a "cooking recipe" (Lucero, Valcke & Schellens, 2013; McLaughlin & MacFadden, 2014). With this approach to ExA's, the students tend to focus their attention on completing the instructions rather than learning about what they are doing and why they are doing it (Hodson, 2005). In other words, the students come to see the method used to carry out the ExA as being more important than what they were expected to learn with it (Bevins & Price, 2016).

Numerous studies also reveal the difficulties that students usually have when facing the different tasks required in an ExA as inquiry, such as posing researchable questions (Hofstein, Navon, Kipnis & Mamlok-Naaman, 2005), formulating hypotheses (Guisasola, Ceberio & Zubimendi, 2006) and understanding the epistemology involved (Kyza, 2009), taking reliable measurements and data (Barolli, Laburú & Guridi, 2010), identifying and controlling variables (Schwichow, Zimmerman, Croker & Härtig, 2016), and interpreting data (Kanari & Millar, 2004), among other aspects.

Even so, the ExA's tend to generate extra motivation in the students to learn science (Bulunuz, 2012), and they are often preferred to other more usual activities in science classes (Abrahams, 2009). However, if the ExA's are posed as a routine, without being stimulating challenges for the learning of science through inquiry, their didactic effectiveness becomes limited and the students may end up losing interest (García-Carmona et al., 2016).

In many cases, the educational ineffectiveness of ExA's is due to the fact that the students do not possess the scientific knowledge and abilities necessary to address them satisfactorily (García-Carmona et al., 2014; Arnold, Kremer & Meyer, 2014). This leads to the students not being able to connect what they observe with the scientific knowledge put into play (Abraham & Millar, 2008). Therefore, it does not make much sense to present a scientific inquiry to the students unless they have previously had the possibility of acquiring the procedural and scientific knowledge bases on which the new knowledge is meant to be constructed (Ferrés, Marbá & Sanmartí, 2015).

In short, it is only when the students become progressively more familiar with ExA's that these promote authentic scientific inquiry. Through the teacher's continued and timely support, these activities can become really useful for learning science (Bertsch, Kapelari & Unterbruner, 2014; Koksal & Berberoglu, 2014).

Strategies for implementing experimental activities as inquiry. In order to integrate the ExA's into a science learning through inquiry framework, some authors (Cañal et al., 2016; Harlen, 2014) suggest that teachers: (i) pose questions that require reasoning, explanations, and reflections, where the students can put their intuitive ideas into play and subject them to analysis; (ii) pay especial attention to helping the students in the use and development of abilities to collect data and interpret evidence; (iii) favour the access to various procedures and ideas through the discussion and handling of different sources of information; (iv) propose challenging tasks, with adequate teacher support, so that the students feel like true researchers; (v) recognize the students' achievements during experimentation, and encourage them to continue until their work culminates in finishing the proposed inquiry; (vi) get the students to properly record their observations during the inquiry, so as to make their subsequent interpretation and discussion of the results easier; (vii) promote the students' discussions about the procedures they used, the results they obtained, etc., in both the small working groups and the class as a whole; and (viii) allow the students time to reflect on how they learnt, what they learnt, and how this can be applied in their future learning about other scientific questions.

When setting ExA's in class, the teacher needs also to bear in mind the types of inquiry that the students can be asked to do. Often it is the character of the question formulated to initiate the scientific inquiry which determines the type of ExA (Harlen, 2013). In this sense, Cañal et al. (2016) propose the following typology: (i) ExA's observing a phenomenon (e.g., "how does … occur?"), (ii) ExA's determining cause and effect relationships (e.g., "how do you think … influences …?"), and (iii) ExA's designing methods or processes to check/observe/measure (e.g., "how would you measure …?").

The nature of the hypothesis formulated to plan the inquiry also determines the type of ExA, although it is also true that it is not always necessary to test a hypothesis in a scientific inquiry (e.g., if it is of an exploratory type). Hypotheses are statements of tentative solutions to problems that guide the development of the inquiry. Guisasola et al. (2006) distinguish three types of hypotheses depending on the nature of the inquiry: (i) descriptive hypotheses, which affirm something which has not yet been confirmed, and are related to factual knowledge (events, phenomena, facts); (ii) explanatory hypotheses, which speculate on the causes of phenomena (explanations, models, theories…), and whose predictions are based on deductive reasoning; and (iii) procedural hypotheses, which refer to knowledge of how to carry out experiments, measurements, or other procedural practices.

In setting up an ExA, different aspects should be taken into account: the students' characteristics and their familiarity with learning through inquiry (Harlen, 2014), the scientific content involved (Newman et al., 2004), the learning objectives pursued (Hodson, 2014), etc. In addition, it is necessary to establish the degree of either openness or guidance with which the activity is going to be implemented in class (Bunterm et al., 2014). Several studies have shown that usually inquiry is more effective when it is more guided, even with students who are familiar with this educational approach (Kirschner, Sweller & Clark, 2006). The reason is that they always need some kind of scaffolding to depend on in inquiry-learning processes (Arnold et al., 2014; Kawalkar & Vijapurkar, 2013). Consequently, ExA's should be set in which there is a balance between the challenge and the support provided by the teacher (Bell, Urhahne, Schanze & Ploetzner, 2010). As a result, some authors consider it useful to have a script to orient the students in their approach to an ExA as guided inquiry (e.g., Bell et al., 2010; García-Carmona et al., 2017; Kipnis & Hofstein, 2008). These scripts would normally include a phase of pre-planning the ExA, some guidelines for its implementation, and issues for the students to reflect on metacognitively. This last aspect, in addition to promoting awareness of their achievements and learning difficulties, favours the development of the students' critical thinking (Seraphin, Philippoff, Kaupp & Vallin, 2012) and their understanding of the nature of scientific inquiry (García-Carmona, 2012).

In view of the above, the proposal by Banchi and Bell (2008) establishing four levels of inquiry according to an increasing order of complexity is interesting. These levels are: (i) confirmation inquiry – the teacher gives the students a question for inquiry which has a previously known answer, and indicates the procedure to follow for its verification; (ii) structured inquiry – the teacher poses the question for inquiry together with the procedure to be followed, and the students must reach a conclusion based on the evidence offered by the data that they collect; (iii) guided inquiry – the teacher poses a question, and the students have to design their own procedures to collect the data and reach a conclusion; and (iv) open inquiry – the students formulate their own inquiry question, design their own procedure, and draw their own conclusions from the data obtained.

Pre-service primary teachers training in inquiry-based science education

The above is also extensible to PPTs. It is interesting, however, to make a brief specific review of the usual difficulties that arise in their training in inquiry-based science teaching strategies.

PPTs always begin their training with ideas on how to teach science which have been constructed mainly from the science classes they themselves received during their schooling (Mellado, Blanco & Ruiz, 1998; García-Carmona et al., 2014). In most cases, this experience was based on traditional educational approaches (García-Carmona & Cruz-Guzmán, 2016), and thus generates internal conflicts in the PPTs that often hinder their assimilation of alternative pedagogical approaches such as those which are inquiry based (Cañal et al., 2011; Crawford, 2007; Kim & Tan, 2011).

The insufficient scientific knowledge and abilities with which PPTs generally access their training in science education also represent a major obstacle to them acquiring the teaching skills they will need to be able to teach science through inquiry (Cortés & Gándara, 2006; Yoon, Joung & Kim, 2012). Moreover, this is a fact that they themselves recognize when asked about it (García-Carmona & Cruz-Guzmán, 2016). In this regard, Newman et al. (2004) insist that harmonizing the learning of science contents and the didactics of science has to be an irreplaceable goal in the formation of PPTs. To this end, they suggest programming courses on strategies of learning-through-inquiry that continuously promote situations in which the PPTs combine their role of science students with that of pre-service teachers.

Another cause that hinders PPTs' assimilating strategies of learning through inquiry is their inadequate understanding of the nature of scientific inquiry itself (Capps & Crawford, 2013; Seung, Park & Jung 2014). This is reflected, for example, in the difficulties they have in fully planning a scientific inquiry on a given question (García-Carmona et al., 2017), in designing valid school experiments to test hypotheses (Yoon, Joung & Kim, 2012), and, in general, in managing to enter into the dynamics of an authentic scientific inquiry (Cortés & Gándara, 2006). However, with an appropriate process of instruction on integrating ExA's into science teaching, PPTs usually gain confidence in how to design and develop them. Above all, they come to recognize ExA's as being suitable resources to integrate theory with scientific practice (Cortés & Gándara, 2006; Ozdem, Ertepinar, Cakiroglu & Erduran, 2013).

Consequently, and as noted at the beginning, in PPT education, the best way to promote ExA's as inquiry is for the students to experiment with them as science learners themselves.

Focus of the Study

In accordance with all of the above, we undertook a project to train PPTs in inquiry-based science learning. After determining the difficulties the PPTs have in their initial planning of a scientific inquiry (García-Carmona et al., 2017), it was proposed to analyse their capabilities to complete an ExA set as a guided inquiry (Banchi & Bell, 2008). Given the characteristics of the PPTs involved in this study (as will be seen later) and the complexity of inquiry-based learning for the uninitiated in this approach (Crawford, 2007; Newman et al., 2004; Seung, Park & Jung 2014), the ExA's chosen were those which represented a manageable challenge for these students. Our aim with this was for the students to experiment with simple scientific inquiries so as to be able to assimilate this educational approach and, at the same time, have examples of useful ExA's for their future as primary education science teachers.

The present study analyses one of these ExA's which consisted in solving a problem of experimental measurement, i.e., one catalogued as being the design of methods for taking measurements (Cañal et al., 2016) and which demands a procedural hypothesis to be posited (Guisasola et al., 2006). The choice of this ExA was due, on the one hand, to the acquisition of measurement skills being a basic objective of the primary education science curriculum, and, on the other, to its being a less frequent type of scientific inquiry, at least according to the literature, although just as important as the other types for the development of students' scientific competence (Caussarieu & Tiberghien, 2017; Hodson, 2014; Kanari & Millar, 2004; Priemer & Hellwig, 2016).

Thus, a qualitative study, descriptive-interpretive in profile, was planned, which was guided by the following two research questions: (1) What abilities do the PPTs present relative to solving a measurement problem through a guided inquiry? (2) What metacognitive conclusions do the PPTs reach after conducting the inquiry?

Methods

Participants and context

In the study, there participated 23 PPTs (21 women and 2 men), with an average age of 20.7 years. They comprised one of the subgroups of the Science Teaching course (9 credits), corresponding to the second year of the Degree in Primary Education in a Spanish university. The participants were a convenience sample, since those selected were accessible at the time of the study.

The course is focused on providing PPTs with basic training to attain an initial development of teaching competencies regarding: the purposes of elementary science education; knowledge of student’s conceptions and difficulties in learning science; the school science curriculum; the selection and sequencing of the content; resources and strategies for teaching and learning science including ExA’s as inquiry; and evaluation.

With respect to their academic profile, most of the participants had accessed the Degree course by an academic itinerary unrelated to science. This indicates that these PPTs had a low academic preference for science, as is common among the students of this undergraduate course in Spain (García-Carmona & Cruz-Guzmán, 2016). Also, when reflecting on the science education they had received, many of them recognize that they do not possess the appropriate scientific training to be able to teach science. They also indicate that their science education had generally fomented just superficial rote learning of the content (García-Carmona et al., 2014), and was therefore far from any approach consonant with inquiry-based science learning.

In their Degree course, during the first year the PPTs study several subjects of science fundamentals (15 credits in total), and these programs include their carrying out ExA's. However, these are usually put to them as a "cooking recipe", along the lines that we explained above. In sum therefore, the participants had had no experience with ExA's as inquiry before beginning the subject of Science Teaching.

In accordance with all of the above, and because there was much other content that had to be addressed in the course's program, the plan designed for training the participants in the use of ExA's as inquiry could not be very ambitious in scope and depth.

Data collection

To answer the two research questions of the present study, a questionnaire was prepared consisting of open items. It was divided into two parts: the first referred to the PPTs' ability to plan and execute a scientific inquiry, and the second to their metacognitive reflections concerning the inquiry they had carried out.

The inquiry question

Luisa, a primary student, was playing with lentils, and suddenly wondered: what would the mass of a lentil be?

Imagine that this problem is put to you. How would you solve it?

Planning and development of the inquiry[1]

1. Formulation of a procedural hypothesis. Explain the strategy with which you all intend to address the resolution of the problem.
2. Method used. Explain the method finally used, justifying the strategy changes made (if any) regarding the initial planning (procedural hypothesis).
3. Data collection. Describe in detail the process followed (measurements, calculations, etc.) to obtain the data.
4. Elaboration of a conclusion. From the data obtained, state your conclusion in response to the problem posed.

Figure 1. *Script of the procedural type of scientific inquiry proposed.*

The first part of the questionnaire (Figure 1) was put to the PPTs as the script for planning and developing an ExA as a guided inquiry, i.e., given a question for inquiry by the instructor, the PPTs had to explain how they intended to address its resolution (procedural hypothesis), how they finally did it, justifying the modifications they had introduced (if any) relative to the initial plan, and what conclusions they obtained.

The second part of the questionnaire (Figure 2) was aimed at determining the PPTs' assessments regarding the difficulties they had encountered, the lessons learnt, and their

1 To carry out the ExA, the PPTs had available various electronic and balance scales with sensitivities ranging from 0.01 to 1 g.

proposed improvements for future experimental measurements similar to those in the ExA. As will be seen later, this would allow the researcher to go more deeply into the analysis and to triangulate information (Oliver-Hoyo & Allen, 2006) with a view to determining the educational effectiveness of the ExA that was implemented.

The responses to the two parts of the questionnaire were collected by the PPTs in a report that they delivered at the end of the ExA. The completion of the ExA required a total of three class sessions (approximately 6 hours). The first session was used for the instructor to present the ExA and for the PPTs to plan the inquiry. The second session was used for the students to carry out the inquiry, and the third to share and discuss the results that they had obtained with the rest of their classmates. The PPTs responded to the metacognition questions out of class time, and these responses were included in the report as well.

Final questions for reflection

Once the inquiry has been completed, answer with sincerity and with arguments the following questions:

1. What difficulties have you encountered with the implementation of the proposed ExA?
2. What do you think you have learnt from the proposed ExA?
3. If, in the future, you are again posed the same or a similar problem about measurement, what modifications would you make to how you resolved it?

Figure 2. *Questions for metacognitive reflection in relation to the scientific inquiry carried out.*

In order to carry out the ExA, the PPTs were organized into groups of 3 or 4 components (7 groups in total), in order to promote an environment of permanent interaction and discussion, and hence of cooperative learning. Following the research line of Kawalkar and Vijapurkar (2013), when any group asked for help or some clarification during the activity, the instructor (the author) gave just the amount of support for them to be able to progress, but always maintaining an environment of challenge.

Data analysis

As mentioned above, a descriptive-interpretive analysis was made of the content of the reports that the groups produced, along the lines of other similar research (e.g., García-Carmona et al., 2017; Nivalainen, Asikainen & Hirvonen, 2013). Due to the open character of the questions, the information was analysed by the researcher (and instructor) in three phases to increasingly refine the coding of this, and therefore to achieve a categorization as representative as possible of the abilities and difficulties of the PPTs with regard to the scientific inquiry.

In the first phase, the researcher made a first proposal to categorize the responses on the basis of the patterns or trends observed. After a certain time (approximately 3 months), the responses were re-analysed thus improving the initial categorization of the responses (second phase). This improvement consisted in some cases in reducing the number of categories by re-grouping the responses, and in others in breaking one category into various subcategories. For example, regarding the PPTs' formulation of the procedural hypothesis, in the first categorization only two types of hypotheses were distinguished (a direct vs. an indirect measurement strategy). But, in the second, the researcher found it pertinent to distinguish two subcategories in one of those original two (the indirect measurement strategy), as will be seen later. This second revision of the categorization affected about 20% of the initial groupings of the responses.

About a month later, the researcher made a third revision of the categorization, making some further refinements (third phase). These affected approximately 10% of the responses, mainly referring to the groups' metacognitive reflections. The result of this process was the final categorization presented in the Results section below.

Finally, in order to meet objectivity criteria, low-inference descriptors were included in the presentation of the results (Seale, 1999), being examples of textual responses by the groups that are representative of the categorizations.

Results

In the following, we shall present the results of the analysis of both the responses the groups gave to the different questions of the ExA script and the metacognitive reflections they made.

Formulation of the procedural hypothesis

The aim with the first question of the script was for the groups to anticipate, in the form of a procedural hypothesis (PH), the strategy they would use to resolve the proposed problem. The results are listed in Table 1.

Table 1
Procedural hypotheses formulated by the groups in the initial planning of the scientific inquiry

Procedural Hypotheses (PH)		Proposing groups*
PH1: direct measurement of the mass of a lentil with a scale.		G1, G2, G6
PH2: indirect measurement of the mass of a lentil.	PH2a: A certain number of lentils are required and the mass is calculated and then the ratio: mass/nº lentils.	G2, G3, G5, G6
	PH2b: A certain mass value is pre-set and lentils are added to the scale until it is reached.	G4, G7

** There are groups that propose both methods of measurement in their procedural hypothesis.*

Two different PHs were distinguished. One of them (PH1), presented by 3 groups, consisted in the direct measurement of the mass of a lentil with a scale. As expressed by one of the groups: "For a more accurate measurement we can use a typical laboratory scale that, thanks to previous experience, we know can measure even the smallest mass."

The other hypothesis (PH2) proposed measuring the mass of a lentil in an indirect way. This hypothesis was presented in two variants. The first, PH2a, was to take a certain number of lentils, measure their mass, and then divide it by the number of lentils to find the mass of a single one (4 groups). An example of how to do this is explained by one of the groups: "… several lentils would be placed on a scale, weighed, and then counted to divide the total mass by the number of lentils, and thus obtain the mass of a single one."

The second variant (PH2b) proposed pre-setting a certain mass value and adding lentils to the scale until reaching that mass, then dividing this by the number of lentils used (2 groups). An example of explaining how to do this is:

> We could measure it with a balance scale in which we put on one saucer a weight of 1 g, and on the other saucer add lentils until the scale is balanced. Then we have to divide 1 g by the number of lentils added to know the mass of the unit.

Table 2
Procedural strategy used in the inquiry

Groups	Method initially planned	Method finally used	Justification for the change of method
G1	PH1	PH2b	The scale does not have sufficient sensitivity to measure the weight of a single lentil, and they decide to change to the method of indirect measurement.
G2	PH1 and PH2a	PH2a	They do not justify why in the end they used only one method, the indirect measurement.
G3	PH2a	PH2a	-
G4	PH2b	PH2a	By adding lentils one by one until a predetermined mass value, they noticed that the scale was not very reliable.
G5	PH2a	PH2a	-
G6	PH1 and PH2a	PH1 and PH2a	-
G7	PH2b	PH1 and PH2b	They finally used the direct and indirect measurement methods to compare the results.

Method of measurement employed

Aware that the procedural hypotheses the groups formulated at the beginning might be modified once they began to make the measurement, we needed to ask them about the strategy they finally used, and the reason for any modification. The results of this question in the script are presented in Table 2.

Three groups maintained their initial procedural hypothesis as the method they finally used to make the measurement. The rest of the groups made some changes to their planned strategy, although one of them did not justify why they had changed their strategy.

The group G1 changed from PH1 to PH2b because they found that the scale they used did not have enough sensitivity to measure the mass of a single lentil:

> The strategy was to observe our hypothesis using a scale and lentils. First we took a lentil and placed it on the scale; we observed that one lentil did not weigh anything, so we poured lentils on the scale until with a limited number of them the scale got a result.

The group G4 always maintained an indirect measurement strategy, although they started with the PH2b strategy and ending up using PH2a. The reason they gave was that with the PH2b strategy they observed that the precision and sensitivity of the scale did not seem reliable to them as it gave "disparate data":

> We used two strategies … (1st) Go on adding lentil by lentil onto the scale dish till reaching 0.1 g, we would count the number of lentils and … get a 'mean' estimate of the weight of a lentil. (2nd) An initial value of lentils to weigh was decided upon (90 in this case) and the same calculation was used as in the previous case … The first strategy gave disparate data. By adding lentils to obtain a value of 0.1 g … what happened was that the first value the scale showed was much higher than 0.1 g … This told us that the scale we were using was not working properly. (…) We went directly to using the 2nd strategy … more satisfactory in this case.

The group G7 went from initially proposing a direct measurement strategy (PH1) to combining this with an indirect measurement, in this case PH2b. The intention was to check whether the two methods arrived at a similar result: "Initially we measured as many lentils as reached 1 g and divided this by the number of lentils … We also weighed one lentil to check the accuracy of the experiment and see if it matched the calculations."

Data collection and conclusion of the inquiry

After this, the groups had to explain in detail in their reports the calculations, measurements, and data obtained to reach a conclusion for the problem posed. The detailed analysis of this issue is presented in Table 3.

Table 3

Data collection and determination of the desired value

	Data collection process					
	One measurement	Several measurements				
		Average of data obtained with one method	Average of data obtained with two methods	Process description	Validation and data reliability	Mass obtained
G1	X			Measure the mass of 130 lentils (7 g) and calculate the mass/nº of lentils ratio.	No	0.050 g
G2		X		They measured 10 lentils with three different scales, one of which gave an inconsistent result. They rejected it and stayed with the results obtained with the other two scales. The result is the average of the two.	No	0.070 g
G3	X			They measured the mass of 118 lentils (7.7 g) and calculated the mass/nº of lentils ratio.	No	0.065 g
G4		X		They noticed the irregular size of the different lentils and decided to make 6 measurements of the mass of 90 lentils, and then averaged the values obtained in each measurement.	Comparison of the data with other groups.	0.064 g
G5		X		They measured the mass of 50 lentils (1.1 g) and calculated the mass/nº of lentils ratio.	No	0.022 g
G6			X	First, they measured the mass of 50 lentils (2 g) and calculated the mass/nº of lentils ratio. Then, with the direct method, they measured the mass of a large lentil, that of a small one, and calculated the average. They obtained a very similar value with each of the two methods.	No	0.040 g
G7			X	With the direct method, they measured the mass of a large lentils, that of a small one, and calculated the average. Then, using the indirect method, they classified the lentils into small (33 lentils ~ 1 g) and large (14 lentils ~ 1 g), measured the masses corresponding to each size, and calculated the average. Finally they took the average of the values obtained by the direct and indirect methods.	Comparison of the data obtained by the direct and indirect methods.	0.050 g

First, we examined whether the groups used one or more measurements to obtain the value of the mass of one lentil. It was noted that two groups only made one measurement to calculate the mass of the lentil, whereas the rest made several measurements following different strategies. In the latter cases, 3 groups were found to have obtained the mass value by calculating the average of the values obtained using a single method. For example, the group G2 measured the mass of 10 lentils with three different scales, one of which they rejected because it gave an incoherent result, thus staying with the results obtained with the other two scales. The end result they proposed was the average of the two measurements (0.070 g).

Another 2 groups combined or compared the results obtained by the direct and indirect methods. Thus, the group G6 first measured the mass of 50 lentils (2 g) and calculated that mass/nº of lentils. Then, using the direct method, they measured the mass of a large lentil and a small one and calculated the average of these. In comparing the results obtained by the two methods, they found that they were very similar. However, with the result not being exactly the same in the two cases, they did not explain whether the solution they were presenting (0.040 g) was obtained by one method, or was the average of the two. The group G7 employed a strategy similar to that of G6, but explained that the final value obtained (0.050 g) was the result of the average values obtained by the direct and indirect methods.

We also analysed whether the groups had employed a strategy of validation and reliability of the results obtained with the inquiry. Only two groups indicated that they had done something in this regard. The group G4 stated that they had compared their result with that obtained by other groups (although they did not explicitly say so, the results indicated that these would have been the groups G2 and G3), and in observing that they were very much alike they were satisfied. Instead, the group G7 took into account the fact that the value of the mass obtained by two different methods (direct and indirect) were very similar (the final result, as mentioned before, was the average of the two measurements).

We noted that each group had a different result for the mass of one lentil, although in some cases (for example, the values obtained by groups G2, G3, and G4) they were quite close to each other. In any case, all the results were in the order of hundredths of a gram.

Metacognitive reflection on the scientific inquiry that had been conducted

The personal assessments of the groups of PPTs on the inquiry they had done (i.e., learning achieved, difficulties encountered, and areas for improvement in future inquiries) allowed the researcher to compare and complete the information (i.e., data triangulation) in order to determine the educational effectiveness of the ExA. In the following we shall present the most interesting results in this regard, with a synthesis given in Table 4.

Table 4
Results of the PPT groups' metacognitive reflections

Metacognitive aspects	Results	Groups
Self-assessment of the learning achieved	Awareness of the importance of making several measurements to refine the results	G3, G7
	Awareness of the importance of choosing the measuring instrument with the most appropriate sensitivity and precision	G5, G7
	Awareness of the complexity of making precise measurements with unsophisticated instruments, similar to the past	G4
	Satisfaction of the curiosity of knowing the mass of a lentil, and motivation for new questions	G2, G5
	Awareness of the need to seek representative measurements of objects or bodies representative of a certain class when these are not identical	G7
	No noteworthy learning is acquired	G1
Self-assessment of the difficulties encountered	Difficulties in the adequate selection of the sample of the elements to measure	G2, G4, G6, G7
	Difficulties in handling the measuring instrument	G3, G5, G6, G7
	Unawareness of the need to validate the result with some known measurement as referent	G1
Proposals to improve future inquiries	More attention to the choice of the most reliable and sensitive instrument possible for the measurement	G2, G3, G7
	Using objects with known measurements as referents against which to validate the results	G3, G5
	Improving the process of selection of the sample of elements to measure, and taking into account a margin of error in the measurement results	G2, G4

Self-assessment of the learning acquired with the inquiry. Only group G1 stated that they had not learnt anything new: "Our experience had no difficulty at all, and it was to carry out an action as common as weighing something; so, in general, we have not learnt anything we did not know …".

The remaining six groups did state having learnt something with the ExA. Thus, groups G3 and G7 referred to the importance of making several measurements to "fine tune" the result. One of them stated: "… we have learnt that to know a figure more accurately when measuring it, we have to do this several times and not be satisfied with the first result."

Groups G2 and G5 highlighted as a positive aspect of the inquiry having satisfied their curiosity about the value of the mass of a lentil and to ask new questions: "… we did not know the actual weight of one lentil, and this made us wonder how many lentils there might be in a plate of lentils that we usually eat."

Groups G5 and G7 highlighted the importance of selecting an instrument with adequate precision and sensitivity for the measurement, as well as properly calibrating the instrument before measuring:

> We learnt that not all scales are useful for the same thing nor do we always need such a specific weight because it is not the same to weigh a backpack (you really do not need to know the decimal part of its weight but just its approximation) as a sharpener (which we need to know its specific weight as it is so small). (G5)
>
> We learnt such aspects as that with electronic scales, the same as with traditional scales, they have to be calibrated. (G7)

Group G4 stressed that the experience was interesting so as to become aware of the complexity of making precise measurements with unsophisticated instruments, referring to the potential difficulties that scientists from the past may have faced: "We became aware of how complex it is to take measurements with traditional scales. The dedication and precision of the experiments by scientists who lacked modern instruments is admirable."

Group G7 also alluded to having become aware that in nature there are no two identical bodies even if they are from the same class (in this case, lentils), and, therefore, it is only possible to give a representative average measurement of the body:

> (…) we have learnt that there is no standard weight for lentils, and in general for many elements found in nature, as each of them has a different size which influences their mass. Therefore, to know the mass of one lentil we can only give a representative figure.

Self-assessment of the difficulties encountered in the inquiry. Four groups (G2, G4, G6, and G7) stated having had difficulties in the selection of the sample of the elements to be measured. One of these groups expressed it as follows:

> (…) the problem that they [the lentils] gave us in their management was the irregularity of their size and therefore mass. In an ideal case in which all were the same, it would have sufficed to make a single measurement, or just a few to make sure they were really the same and that there had been no mistake in measuring. But in reality, the disparate size led us to debate between two ways of acting when weighing each set of 90 lentils.

The difficulty in handling the measuring instrument (scale) was also referred to by four groups (G3, G5, G6, and G7). One of the groups stated it this way in their final report: "We did not know how to use the first scale that we took, and also we had not calibrated the digital scale because we did not know that this had to be done."

Also, a difficulty that group G1 highlighted as a mistake they had made during the scientific inquiry was not applying any criterion to validate their result which proved to be inconsistent. However, they became aware of this, and resolved it with the aid of the instructor:

> (…) when we got the result that the lentil had a mass of 2 g, we were pleased with this result without thinking of its coherence. (…) The instructor told us to think about a paracetamol pill of 1 g … How would just one lentil weigh 2 g? Immediately we carried out the experiment again … accepting that taking the 2 g as being the correct answer was a mistake.

Proposals to improve the inquiry. Three groups (G2, G3, and G7) indicated that, if they did the inquiry again, they would pay more attention to selecting a more reliable or more sensitive measuring instrument (scale). An example of a response in this regard: "We think we would make changes in the planning of the materials, and would eliminate the idea of using a balance scale and a gram weight, to use directly the digital scale."

Groups G3 and G5 emphasized that they used objects with known masses as referents to validate the results. One of them explained it like this: "… we would be concerned about having other objects as reference, thus getting approximate values and reducing the margin of error. For example, if a chickpea weighs '5', a lentil can not weigh ‘10’."

Groups G2 and G4 would improve the measurement strategy used, focusing on a better selection of the sample of lentils and the need to consider a margin of error in the result. One of these groups expressed it as follows:

> We would put all the lentils through the different sieves available, grouping them depending on their size. With this we could weigh average-sized lentils or weigh samples from each group and then take the average. The advantage of prior sieving is that we would have the necessary data to say in what percentage we would find the mean weight of a lentil or give the final result with a margin of error of $\pm x$ grams for the whole sample.

Discussion

The results indicate that, overall, the PPTs were able to plan and complete a scientific inquiry with a quite acceptable degree of satisfaction. This is remarkable considering the usual academic profile of the PPTs, characterized by limited scientific training and no experience with the inquiry-based educational approach (García-Carmona & Cruz-Guzmán, 2016; Newman et al., 2004).

The result might be a consequence of the nature of the proposed inquiry question which demanded a procedural resolution unconditioned by any specific scientific knowledge. This idea is reinforced by the fact that when PPTs are faced with another kind of scientific inquiry, the results are often less satisfactory. Indeed, PPTs have considerable difficulty in planning a scientific inquiry of a cause and effect type because this requires them to formulate an explanatory hypothesis (García-Carmona et al., 2017; Yoon, Joung & Kim, 2012), and this has to be based on scientific knowledge that PPTs generally have not sufficiently acquired. This is a common difficulty even in the training of pre-service science teachers of other educational levels (Aydoğdu, 2015; Nivalainen, Asikainen, Sormunen & Hirvonen, 2010), despite their having greater scientific knowledge.

Therefore, given the usual difficulties that PPTs have in assimilating such a complex educational approach as that based on inquiry (Crawford, 2007; Seung, Park & Jung 2014), a proposal of an ExA such as that analysed here may be an ideal way of introducing them to this focus. According to Kim and Tan (2011), PPTs often begin their training in science teaching very reluctant to consider inquiry-type ExA's as part of their teaching plans. Thus, any attempt to start their training in this educational approach with more ambitious ExA's may end up being frustrating for them, and consequently discourage them from promoting ExA's in their future science classes (Cañal et al., 2011).

That said, it is important to separate out the PPTs' strengths and weaknesses regarding the ExA analysed in the present study. Regarding the strengths, there stands out firstly the variety of procedural hypotheses (i.e., two different indirect measurement strategies and one direct) that the groups initially formulated in planning the inquiry. Also, each group had a specific method of calculation that was different from the others (see Table 3). This demonstrates the usefulness of the proposed ExA to develop students' creative and divergent thinking in an inquiry-based learning context (Yang, Lee, Hong & Lin, 2016).

Secondly, several groups re-thought their initial hypotheses, adapting them to the limitations and/or difficulties they encountered during the measurement process. I.e., they were aware of the need to change the initial hypothesis they had posited so as to adapt it to the requirements that came up during the inquiry (Sadeh & Zion, 2009). This contrasts with what is the case when students are faced with cause-effect scientific inquiries, in which they show themselves to have difficulties in understanding the role of alternative hypotheses when the evidence they obtain does not support their initial hypothesis (Kyza, 2009).

Regarding the weaknesses, the most notable ones were related to data collection and getting the final result. Most of the groups did not use any validation and reliability strategy. Only two groups commented on having done something in this respect, although with strategies that were not really scientifically rigorous. In addition, 2 of the 7 groups were content to make only one measurement to obtain the final result of the inquiry, and none of them considered a margin of measurement error in their presentation of the result. They therefore showed a deficiency in these skills that is common among students who are faced with having to make empirical measurements (Barolli et al., 2010; Kanari & Millar, 2004; Priemer & Hellwig, 2016). Consequently, the abilities needed in making measurements should receive special attention in training PPTs (Godino, Batanero & Roa, 2002). Nevertheless, the PPTs' final metacognitive reflections regarding these aspects suggests that they can acquire these basic skills without too much effort in their training as pre-service teachers.

Indeed, in the PPTs' self-assessment of their learning and difficulties in the inquiry they had carried out, it was mentioned that, in an empirical measurement, the data has to be acquired with several measurements and using suitable and reliable instruments. In addition, some groups referred to the need to have objects with known values of the measurement as references against which to validate the result. There were even some cases which cited the need to consider a margin of error in the final experimental result. All this showed the usefulness of the ExA implemented to improve the understanding of certain aspects of the nature of scientific inquiry (Caussarieu & Tiberghien, 2017; García-Carmona, 2012).

Other interesting aspects highlighted by the PPTs in their metacognitive reflections were: the interest of the inquiry to satisfy curiosity (i.e., knowing the approximate mass of a lentil); the recognition of the need to provide measurements that are representative of a magnitude when referring to objects that, although of the same class, are not identical; and becoming aware of the difficulty of making certain measurements when the most appropriate instruments are not available.

Within these reflections, there also stood out the role of the instructor as a facilitator of learning, providing the groups with the appropriate scaffolding for them to progress adequately. In most cases, the instructor did this by formulating questions for the groups to re-think those aspects in which they found difficulties or which had given inconsistent results. As in other studies (Kawalkar & Vijapurkar, 2013), this strategy proved effective as was indicated by one group who noted that, thanks to a question put by the instructor, they realized that the result they had obtained for the mass of a lentil was incongruent, and therefore that they had to revise their process.

Consequently, the PPTs' metacognition process was especially useful, firstly because it allowed the researcher to go into the analysis in greater depth and corroborate the findings on the educational effectiveness of the ExA (triangulation of the information); secondly because it encouraged the PPTs' critical thinking (Seraphin et al., 2012) in accepting their mistakes and recognizing having learnt from them (Kipnis & Hofstein, 2008); and thirdly because it favoured the understanding of some aspects of the nature of science, such as the possibility of reaching the same result through different procedures, or the importance of refining the experiments to obtain more reliable results, both being aspects that are often difficult for PPTs to assimilate (García-Carmona & Acevedo, 2016).

Implications, Limitations, and Perspectives

The insufficient scientific background that PPTs usually have, their lack of familiarity with inquiry-based learning, and the complexity of that approach are all factors that hinder the formation of PPTs in this line. Consequently, it is necessary for the instructors of PPTs to seek new ways to help them meet this challenge "without dying in the attempt". In this regard, the main contribution of this study has been to show that, with procedural type ExA's such as that which we have described, the PPTs become involved in an authentic scientific inquiry with little difficulty. This contrasts with the less satisfactory results usually obtained when PPTs are faced with an explanatory ExA (i.e., one of cause and effect) because this requires the use of scientific knowledge that the PPTs often do not possess adequately.

However, with the above we do not mean to say that the training of PPTs should include only procedural ExA's and other types should be discarded. What we suggest is the combination of different types of ExA. And not only for training in the inquiry-based learning approach to be fuller, but also because carrying out some types of ExA may help improve the educational effectiveness of others. For example, a procedural type ExA may help PPTs improve their skills in handling measuring instruments, in collecting and processing data, etc., which are also essential tasks in other types of ExA (Kanari & Millar, 2004), as well as contributing to their understanding of the nature of scientific inquiry if the corresponding reflections are promoted (García-Carmona, 2012).

In sum, for the training of PPTs in inquiry-based science learning by means of ExA's, we would suggest that instructors prepare catalogues of different types of simple ExA's, and implement them throughout the course in an order of increasing difficulty as the PPTs acquire the essential skills with which to address a scientific inquiry.

Finally, it is necessary to refer to the limitations and perspectives of the study we have presented. The results come from a specific context of PPT training, and we have only analysed the educational effectiveness of a procedural ExA on a specific measurement (the mass of a lentil). The conclusions should therefore be taken in the educational context in which the research was conducted. Nonetheless, these findings may be useful to stimulate further research on the issue that was addressed in educational contexts that are similar to that described here. Above all, they may be useful to promote new ideas and approaches about how to effectively train PPTs in inquiry-based science teaching. Thus, it would be interesting to analyse in the future the effectiveness of other ExA's on measuring different types of magnitudes, both fundamental (mass, time, temperature…) and derivative (volume, pressure, density…) in the line proposed here. This would allow one to obtain more robust results to determine whether this type of ExA is indeed conducive to engaging the PPTs in genuine scientific inquiry. It would also be recommendable to analyse to what extent the learning acquired with procedural ExA's favour the PPTs' effective performance of other ExA's, such as those that investigate cause-effect relationships.

Funding Information This work was supported by the Ministry of Economy and Competitiveness (Spain) under grant EDU2013-41003-P.

References

Abd-El-Khalick, F., Boujaoude, S., Duschl, R., Lederman, N. G., Mamlok-Naaman, R., Hofstein, A., . . . Tuan, H.-L. (2004). Inquiry in science education: International perspectives. Science Education, 88(3), 397–419.

Abrahams, I. (2009). Does practical work really motivate? A study of the affective value of practical work in secondary school science. International Journal of Science Education, 31(17), 2335–2353.

Abrahams, I., & Millar, R. (2008). Does practical work really work? A study of the effectiveness of practical work as a teaching and learning method in school science. International Journal of Science Education, 30(14), 1945–1969.

Arnold, J. C., Kremer, K., & Mayer, J. (2014). Understanding students' experiments—What kind of support do they need in inquiry tasks? International Journal of Science Education, 36(16), 2719–2749.

Aydoğdu, B. (2015). Examining preservice science teachers' skills of formulating hypotheses and identifying variables. Asia-Pacific Forum on Science Learning and Teaching, 16(1), 1–38.

Banchi, H., & Bell, R. (2008). The many levels of inquiry. Science and Children, 46(2), 26–29.

Barolli, E., Laburú, C. E., & Guridi, V. M. (2010). Laboratorio didáctico de ciencias: Caminos de investigación [Didactic laboratory of science: Research paths]. Revista Electrónica de Enseñanza de las Ciencias, 9(1), 88–110.

Bell, T., Urhahne, D., Schanze, S., & Ploetzner, R. (2010). Collaborative inquiry learning: Models, tools and challenges. International Journal of Science Education, 32(3), 349–377.

Bertsch, C., Kapelari, S., & Unterbruner, U. (2014). From cookbook experiments to inquiry based primary science: Influence of inquiry based lessons on interest and conceptual understanding. Inquiry in Primary Science Education, 1, 20–31.

Bevins, S., & Price, G. (2016). Reconceptualising inquiry in science education. International Journal of Science Education, 38(1), 17–29.

Bulunuz, M. (2012). Motivational qualities of hands-on science activities for Turkish preservice kindergarten teachers. Eurasia Journal of Mathematics, Science & Technology Education, 8(2), 73–82.

Bunterm, T., Lee, K., Lan, J. N., Srikoon, S., Vangpoomyai, P., Rattanavongsa, J., & Rachahoon, G. (2014). Do different levels of inquiry lead to different learning outcomes? A comparison between guided and structured inquiry. International Journal of Science Education, 36(12), 1937–1959.

Cañal, P., Travé, G., & Pozuelos, F. J. (2011). Análisis de obstáculos y dificultades de profesores y estudiantes en la utilización de enfoques de investigación escolar [An analysis of the teachers' and students' obstacles and difficulties on the use of inquiry approach]. Investigación en la Escuela, 73, 5–26.

Cañal, P., Criado, A. M., García-Carmona, A., & Muñoz, G. (2013). La enseñanza relativa al medio en las aulas españolas de Educación Infantil y Primaria: Concepciones didácticas y práctica docente [The teaching related to the environment in the Spanish classrooms of Infant and Primary Education: Didactic conceptions and teaching practice]. Investigación en la Escuela, 81, 21–42.

Cañal, P., García-Carmona, A., & Cruz-Guzmán, M. (2016). Didáctica de las Ciencias Experimentales en Educación Primaria [Didactics of Experimental Sciences in Primary Education]. Madrid, Spain: Paraninfo.

Capps, D. K., & Crawford, B. A. (2013). Inquiry-based professional development: What does it take to support teachers in learning about inquiry and nature of science? International Journal of Science Education, 35(12), 1947–1978.

Caussarieu, A., & Tiberghien, A. (2017). When and why are the values of physical quantities expressed with uncertainties? A case study of a physics undergraduate laboratory course. International Journal of Science and Mathematics Education, 15(6), 997–1015.

Cortés, A. L., & Gándara, M. (2006). La construcción de problemas en el laboratorio durante la formación del profesorado: Una experiencia didáctica [Building problems in the laboratory during teacher training: A didactic experience]. Enseñanza de las Ciencias, 25(3), 435–450.

Cortés, A. L., Gándara, M., Calvo, J. M., Martínez, M. B., Ibarra, M., Arlegui, J., & Gil, M. J. (2012). Expectativas, necesidades y oportunidades de los maestros en formación ante la enseñanza de las ciencias en la educación primaria [Expectations, needs and opportunities of pre-service teachers in view of science teaching in primary education]. Enseñanza de las Ciencias, 30(3), 155–176.

Crawford, B. A. (2007). Learning to teach science as inquiry in the rough and tumble of practice. Journal of Research in Science Teaching, 44(4), 613–642.

Criado, A. M., & García-Carmona, A. (2011). Las experiencias prácticas para el conocimiento del medio (natural y tecnológico) en la formación inicial de maestros [The practical experiences for the knowledge of the environment (natural and technological) in the initial formation of teachers]. Investigación en la Escuela, 74,73–88.

Cruz-Guzmán, M., García-Carmona, A. & Criado, A. M. (2017). An analysis of the questions proposed by elementary pre-service teachers when designing experimental activities as inquiry. International Journal of Science Education. https://doi.org/10.1080/09500693.2017.1351649.

Ferrés, C., Marbà, A., & Sanmartí, N. (2015). Trabajos de indagación de los alumnos: Instrumentos de evaluación e identificación de dificultades [Students' inquiry works: Assessment tools and identification of difficulties]. Revista Eureka sobre Enseñanza y Divulgación de las Ciencias, 12(1), 22–37.

Flores, J., Caballero, M. C., & Moreira, M. A. (2009). El laboratorio en la enseñanza de las ciencias: Una visión integral en este complejo ambiente de aprendizaje [The science laboratory teaching: An integral vision in this complex learning environment]. Revista de Investigación, 33, 75–111.

García-Carmona, A. (2012). Cómo enseñar naturaleza de la ciencia (NDC) a través de experiencias escolares de investigación científica [How to teach Nature of Science (NDC) through scholarly scientific research experiences]. Alambique, 72, 55–63.

García-Carmona, A., & Acevedo, J. A. (2016). Concepciones de estudiantes de profesorado de Educación Primaria sobre la naturaleza de la ciencia: Una evaluación diagnóstica a partir de reflexiones en equipo [Concepts of primary school teacher students on the nature of science: A diagnostic evaluation based on team reflections]. Revista Mexicana de Investigación Educativa, 21(69), 583–610.

García-Carmona, A., & Cruz-Guzmán, M. (2016). ¿Con qué vivencias, potencialidades y predisposiciones inician los futuros docentes de Educación Primaria su formación en la enseñanza de la ciencia? [What experiences, potentials and predispositions initiate future teachers of Primary Education training in teaching science?]. Revista Eureka sobre Enseñanza y Divulgación de las Ciencias, 13(2), 440–458.

García-Carmona, A., Cruz-Guzmán, M., & Criado, A. M. (2014). "¿Qué hacías para aprobar los exámenes de ciencias, qué aprendiste y qué cambiarías?". Preguntamos a futuros docentes de Educación Primaria ["What did you do to pass the science exams, what did you learn and what would you change?" We ask future teachers of Primary Education]. InvestigaciónenlaEscuela, 84, 31–46.

García-Carmona, A., Criado, A. M., & Cruz-Guzmán, M. (2016). Prospective primary teachers' prior experiences, conceptions, and pedagogical valuations of experimental activities in science education. International Journal of Science and Mathematics Education. https://doi.org/10.1007/s10763-016-9773-3.

García-Carmona, A., Criado, A. M., & Cruz-Guzmán, M. (2017). Primary pre-service teachers' skills in planning a guided scientific inquiry. Research in Science Education, 47(5), 989–1010.

Godino, J. D., Batanero, M. C., & Roa, R. (2002). Medida de magnitudes y su didáctica para maestros [Elementary teacher education on measurement of magnitudes]. Granada, Spain: Universidad de Granada.

Guisasola, J., Ceberio, M., & Zubimendi, J. L. (2006). University students' strategies for constructing hypothesis when tackling paper-and-pencil tasks in physics. Research in Science Education, 36(3), 163–186.

Harlen,W.(2013).Assessment & inquiry-based science education: Issues in policy and practice. Trieste, Italy: IAP.

Harlen, W. (2014). Helping children's development of inquiry skills. Inquiry in Primary Science Education, 1, 5–19.

Hodson, D. (2005). Teaching and learning chemistry in the laboratory: A critical look at the research. Educación Química, 16(1), 30–38.

Hodson, D. (2014). Learning science, learning about science, doing science: Different goals demand different learning methods. International Journal of Science Education, 36(15), 2534–2553.

Hofstein, A., Navon, O., Kipnis, M., & Mamlok-Naaman, R. (2005). Developing students' ability to ask more and better questions resulting from inquiry-type chemistry laboratories. Journal of Research in Science Teaching, 42(7), 791–806.

Kanari, Z., & Millar, R. (2004). Reasoning from data: How students collect and interpret data in science investigations. Journal of Research in Science Teaching, 41(7), 748–769.

Kawalkar, A., & Vijapurkar, J. (2013). Scaffolding science talk: The role of teachers' questions in the inquiry classroom. International Journal of Science Education, 35(12), 2004–2027.

Kim, M., & Tan, A.-L. (2011). Rethinking difficulties of teaching inquiry-based practical work: Stories from elementary pre-service teachers. International Journal of Science Education, 33(4), 465–486.

Kipnis, M., & Hofstein, A. (2008). The inquiry laboratory as a source for development of metacognitive skills. International Journal of Science and Mathematics Education, 6(3), 601–627.

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. Educational Psychologist, 41(2), 75–86.

Koksal, E. A., & Berberoglu, G. (2014). The effect of guided-inquiry instruction on 6th grade Turkish students' achievement, science process skills, and attitudes toward science. International Journal of Science Education, 36(1), 66–78.

Kyza, E. A. (2009). Middle-school students' reasoning about alternative hypotheses in a scaffolded, softwarebased inquiry investigation. Cognition and Instruction, 27(4), 277–311.

Lucero, M., Valcke, M., & Schellens, T. (2013). Teachers' beliefs and self-reported use of inquiry in science education in public primary schools. International Journal of Science Education, 35(8), 1407–1423.

McLaughlin, C. A., & MacFadden, B. J. (2014). At the elbows of scientists: Shaping science teachers' conceptions and enactment of inquiry-based instruction. Research in Science Education, 44(6), 927–947.

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

Minner, D. D., Levy, A. J., & Century, J. (2010). Inquiry-based science instruction—What is it and does it matter? Results from a research synthesis years 1984 to 2002. Journal of Research in Science Teaching, 47(4), 474–496.

Newman, W. J., Abell, S. K., Hubbard, P. D., McDonald, J., Otaala, J., & Martini, M. (2004). Dilemmas of teaching inquiry in elementary science methods. Journal of Science Teacher Education, 15(4), 257–279.

Nivalainen, V., Asikainen, M. A., & Hirvonen, P. E. (2013). Preservice teachers' objectives and their experience of practical work. Physical Review Special Topics—Physics Education Research, 9(1), 010102.

Nivalainen, V., Asikainen, M. A., Sormunen, K., & Hirvonen, P. E. (2010). Preservice and inservice teachers' challenges in the planning of practical work in physics. Journal of Science Teacher Education, 21(4), 393–409.

Oliver-Hoyo, M., & Allen, D. (2006). The use of triangulation methods in qualitative educational research. Journal of College Science Teaching, 35(4), 42–47.

Ozdem, Y., Ertepinar, H., Cakiroglu, J., & Erduran, S. (2013). The nature of pre-service science teachers' argumentation in inquiry-oriented laboratory context. International Journal of Science Education, 35(15), 2559–2586.

Priemer, B. & Hellwig, J. (2016). Learning about measurement uncertainties in secondary education: A model of the subject matter. International Journal of Science and Mathematics Education,1–24. https://doi. org/10.1007/s10763-016-9768-0

Rocard, M., Csermely, P., Jorde, D., Lenzen, D., Walberg, H., & Hemmo, V. (2007). Science education now: A renewed pedagogy for the future of Europe. Brussels, Belgium: Directorate General for Research, Science, Economy and Society.

Sadeh, I., & Zion, M. (2009). The development of dynamic inquiry performances within an open inquiry setting: A comparison to guided inquiry setting. Journal of Research in Science Teaching, 46(10), 1137–1160.

Schwichow, M., Zimmerman, C., Croker, S., & Härtig, H. (2016). What students learn from hands-on activities. Journal of Research in Science Teaching, 53(7), 980–1002.

Seale, C. (1999). The quality of qualitative research. London, England: SAGE.

Seraphin, K. D., Philippoff, J., Kaupp, L., & Vallin, L. M. (2012). Metacognition as means to increase the effectiveness of inquiry-based science education. Science Education International, 23(4), 366–382.

Seung, E., Park, S., & Jung, J. (2014). Exploring preservice elementary teachers' understanding of the essential features of inquiry-based science teaching using evidence-based reflection. Research in Science Education, 44(4), 507–529.

Yang, K. K., Lee, L., Hong, Z. R., & Lin, H. S. (2016). Investigation of effective strategies for developing creative science thinking. International Journal of Science Education, 38(13), 2133–2151.

Yoon, H. G., Joung, Y. J., & Kim, M. (2012). The challenges of science inquiry teaching for pre-service teachers in elementary classrooms: Difficulties on and under the scene. Research in Science Education, 42(3), 589–608.