A five-level design for evaluating professional development programs: Teaching and learning about nature of science

Eda Erdas Kartal
Kastamonu University, Turkey
Nihal Dogan
Abant Izzet Baysal University, Turkey
Serhat Irez
Marmara University, Turkey
Gultekin Cakmakci and Yalcin Yalaki
Hacettepe University, Turkey

The aim of this study is to evaluate a continuing professional development (CPD) program designed to improve teachers' professional competences about the nature of science (NOS) by using a new evaluation perspective. The researchers followed a year long CPD about NOS with the voluntary attendance of 18 middle school science teachers and their students. In a collaborative and reflective environment, teachers were introduced to various NOS aspects and ways to use explicit instruction and formative assessment in their NOS teaching. In addition, teachers received teaching activities and materials to be implemented in their classrooms for one year. The data were collected and evaluated based on the "five level (learning, beliefs, transfer, results, and reaction) evaluation model". The findings demonstrate that the CPD program about NOS effectively improved teachers' views about NOS, beliefs about teaching and learning the NOS, classroom practices about NOS and also their students' views about NOS. This study is the first to use formative assessment and discourse analysis in a professional development program for in-service teacher education, and also first to evaluate teachers' views, beliefs, practices and their students' views about NOS all together. The findings are thought to motivate researchers to consider multiple level evaluations of future professional development programs.

Introduction

The nature of science (NOS), generally defined as "the values and assumptions inherent in the evolution of scientific knowledge", is an important component of scientific literacy. To achieve the vision of scientific literacy, teachers should have professional competence about NOS. Ongoing effective professional development opportunities are important venues for helping teachers improve their professional competence in the NOS (Akerson & Hanuscin, 2007).

Professional development is defined as the provision of activities designed to improve teachers' knowledge, skills, and understanding, leading to changes in their thinking and classroom behaviour (Fenstermacher & Berliner, 1983). Professional development can be a critical factor for teacher development and student achievement (Lowden, 2005). Specifically, reforms aimed at improving science teaching and learning can be achieved when teachers receive the appropriate professional development (Dass & Yager, 2009). Teacher professional development has undergone changes over the years. As teachers' professional development became more prominent in science education reform, it has also became more obvious that the existing traditional forms of professional development were not sufficiently successful. Traditional professional development, which generally included quick solutions that were compressed into an after-lunch or a single full-day session, received mounting criticism from teacher education researchers (Fullan, 1995).

Thus, a new perspective arose and various changes have been made to the traditional professional development system (Renyi, 1996). As part of these changes, curriculum-based, effective continuing professional development (CPD) programs are now often preferred to curriculum-independent short-term professional development. We can describe an effective CPD program as the provision of a learning process resulting from meaningful interaction with the context that eventually leads to changes in teachers' knowledge, practices, and in their thinking about those practices (Kelchtermans, 2004). Research shows that teachers' participation in effective CPD programs improves their views about the NOS (Burton 2013), and classroom practices of the NOS (Akerson & Hanuscin, 2007). In addition, studies demonstrate that these improvements are reflected positively in the learning outcomes of students (Lederman, Lederman, Kim & Ko, 2012). Thus, multiple studies in recent years have focused on the development of teachers' knowledge and practices of NOS through CPD programs (Burton, 2013; Capps & Crawford, 2013).

Researchers have studied extensively the impact of professional development programs on NOS views and practices. In one study, Lederman et al. (2012) looked at the effects of a 5-year professional development program on teachers' and students' views about the NOS and scientific inquiry, and the teachers' pedagogical content knowledge of their practices on the teaching of the NOS and scientific inquiry. Bell and Maeng (2013) explored a research-based professional development program's impact on the views and confidences of teachers, science coordinators, and science educators toward problem-based learning, NOS, and the application of inquiry-based teaching. Burton (2013) probed the effects of a 5-year professional development program on teachers' views and pedagogical content knowledge of NOS. The research findings in this field generally show that teachers' views and practice of the NOS can be improved by providing adequate time and resources as well as by using effective teaching approaches, giving teachers reflection time during the CPD process, and receiving feedback on the CPD program.

The studies mentioned above indicate that the context, features, and implementation methods of CPD programs are very important. However, relying on popular and successful approaches does not guarantee that a CPD program will be effective, because teachers' classroom practices are influenced also by other variables. One of these variables is beliefs that teachers have about teaching and learning. Researchers have argued that these beliefs can influence teachers' teaching strategies and performance in the classroom (Cheng, Chan, Tang & Cheng, 2009). Similarly, research suggests that these beliefs may cause teachers to resist the targeted change in classroom practice (Brickhouse, 1990). Thus, targeting teachers to change their beliefs about teaching and learning is important to ensure that their classroom practices are differentiated and do adapt to change. Therefore, one of the most important objectives of professional development programs, as well as teacher training programs, should be the development of beliefs about teaching and learning, which are formed in the light of past experiences as well as emotional accumulations (Nespor, 1987). Research shows that professional development programs prepared with consideration of teachers' belief systems develop teacher beliefs; it shows that this development increases the effectiveness of classroom practices and student achievement (Lumpe, Czerniak, Haney & Beltyukova, 2012).

One of the teachers' beliefs about teaching and learning that affect their classroom practice is their self-efficacy beliefs (Erdas, 2015). Research shows that it is essential to develop teachers' self-efficacy beliefs in order to effectively adapt the proposed teaching practices to classroom practices (Luft & Hewson, 2014). Another teacher belief about teaching and learning that affect their classroom practices is their belief about new approaches in reform documents, where learning about NOS is emphasised. Although teachers try to adapt to new approaches in classroom practice, the contradiction between their beliefs and practices towards these approaches may cause them to change the new curriculum to the traditional approach (Mitchener & Anderson, 1989). Therefore, developing teachers' beliefs about reform approaches is important in implementing new approaches emphasised in professional development programs.

Research shows that teachers' participation in effective CPD programs improves teacher beliefs about teaching and learning NOS (Bell & Maeng, 2013). In the related literature, some research has aimed to develop teachers' views on the NOS and classroom practices through professional development programs. However, these studies are limited (Lederman et al., 2012). Also there is little research aimed at developing teachers' views about the NOS, their beliefs about teaching and learning of this subject, and their classroom practices concurrently. In addition, researchers generally do not focus on evaluation of the programs they organise, for reasons such as complexity of the evaluation, and the lack of academic consensus on criteria for measuring the effectiveness of professional development programs (Guskey, 2000). But researchers who are responsible for the planning and implementation of professional development programs about this subject should know how to critically assess and evaluate the effectiveness of their actions (Guskey & Yoon, 2009). Some researchers who agree on this issue have suggested various evaluation models in order to evaluate these professional development programs.

One of these models is *Tyler's evaluation model.* This model, developed by Tyler (1942), suggests that setting goals is the most important step in evaluating a professional development program. According to this model, evaluation should focus on the extent to which these goals have been achieved (Guskey, 2000).

Another model is *Scriven's goal-free evaluation model.* According to Guskey (2000), focusing on goals that are not well defined in the evaluation process limits proper evaluation and may cause other products in the program to be neglected. The goal-free model developed by Scriven (1967) is designed to overcome this limitation in evaluation (Kanadli, 2012).

Metfessel and Michael's evaluation model (Metfessel & Michael, 1967) consists of the following eight evaluation steps: including all the members of the school (students, administrators and families) in the evaluation process, developing general and specific goals, transforming goals into a viable form, developing measurement criteria and tools, measuring the degree of attainment of goals, analysing data, and interpreting the data in the light of the established standards, and preparing suggestions for the future by eyeing goals (Kanadlı, 2012).

In *Stufflebeam's evaluation model* (Stufflebeam, 1969), evaluation is a cyclical process. According to this model, evaluation should provide a continuous flow of information to decision makers regarding context, input, process, and product evaluation (Guskey, 2000).

Another model, *Hammond's evaluation model* (Hammond, 1973) argues that researchers should state how the evaluation process was achieved and also the reasons why some targets were not reached while other targets were reached during the evaluation process. Hammond suggested a three-dimensional (cube) model to organise the factors that affect any achievement in this model (Worten & Sanders, 1987): teaching (content, method, tools, and cost), institution (students, teachers, administrators, educational experts, family, and community) and behaviour (cognitive, affective, and psychomotor).

Kirkpatrick's evaluation model (Kirkpatrick, 1959; 2001) contends that various levels of data should be examined to assess the effectiveness of a professional development program. This model was developed to determine the effectiveness of business and industry training programs and was also applied by Guskey (2000) to assess professional development in the field of education. Kirkpatrick summarised the four levels of evaluation in this model (Kirkpatrick, 1959; Kirkpatrick & Kirkpatrick, 2007): reactions (teachers' reactions to the professional development program), learning (teachers' learning about content), transfer (whether the professional development program made a difference in teachers' professional practice) and results (learning outcomes of students). Assessing participants' reactions is the most common and easiest way to evaluate professional development. A positive response from the teachers is a prerequisite for the positive evaluation results. This type of data helps also to improve professional development program design (Guskey, 2000). It should also be noted in effective professional development programs, both teachers' content knowledge and pedagogical knowledge should be developed at the same time (Guskey, 2003). Positive findings about teachers' learning do not mean that their practices have improved as well. For this reason, it is not surprising that the program must be evaluated at these two levels (learning and transfer). Finally, the underlying purpose of professional development programs is to improve student learning. For this reason, as Guskey (2000) points out, measuring student learning outcomes indicates the bottom line in education.

Guskey (2000) stated that Kirkpatrick's four-level evaluation model as used in education had insufficient explanatory power. In order to overcome this deficiency, Guskey (2000) introduced his *five-level evaluation model* which adapted Kirkpatrick's model to the field of education: participants' reactions, participants' learning, organisation support and change, participants' use of new knowledge and skills, and student learning outcomes. Guskey

(2000) stated that organisational support for professional development programs and the realisation of change in the organisations involved could affect the effectiveness of professional development programs. For this reason, Guskey added organisational support and change level to his model.

The five level evaluation model

As Guskey (2000) maintained, evaluating professional development program effectiveness requires data at various levels and tracking efficiency at a limited level provides insufficient information. Evaluating program effectiveness throughout the process provides valuable information to researchers about whether things are going as planned and if the expected progress is being made. If the expected progress is not being made, then evaluative comments can be used to improve the format and organisation of the professional development program. However, as seen in the studies mentioned above, the effectiveness of CPD programs is usually evaluated at one or two levels. Studies evaluating CPD program effectiveness at multiple levels are very limited. Thus, this study relies on a *'five level evaluation model'* to assess the effectiveness of a CPD program about NOS.

The *'five level evaluation model'* implemented in this study is different from Guskey's (2000) five level evaluation model. This study's model does not cover Guskey's (2000) *organisation support end change* evaluation level. The reason for not using Guskey's model is that the focus of this research is on the changes that can be achieved in the teacher and teacher's practices through a CPD program, and its reflection on learning outcomes. This research sheds light on the levels to which the effectiveness of CPD can be evaluated with reference to teacher and student level when the appropriate organisational conditions are met. This model is based on research revealing that CPD programs should be related to the beliefs of teachers. These studies emphasise that it is important to identify, develop and evaluate teacher beliefs in CPD in order to be effective (Haney & Lumpe, 1995; Posnanski, 2002). Thus, the *belief level* has been added to Kirkpatrick's model, and the evaluation of CPD programs at these five levels has been suggested by researchers: reactions (teachers' reactions to the professional development program), learning (teachers' learning about content), belief (whether professional development programs make a difference in teacher beliefs about teaching and learning content), transfer (whether professional development programs make a difference in teachers' professional practice) and results (learning outcomes of students) (Figure 1).

As shown in the Figure 1, the level of belief is more inward level than the levels of learning and reaction, so it is more difficult to change and to evaluate changement. On the other hand, the level of response is the easiest to change and to evaluate change, so it is put on the outermost ring in the model. The level of transfer is a bridge between teacher level development and student level development. Therefore, it is important to evaluate development at transfer level according to this model in terms of explaining development at student level. It is also important to evaluate the level of result, which is an important indicator of development at the student level, in terms of supporting the development in the transfer level (Figure 1).

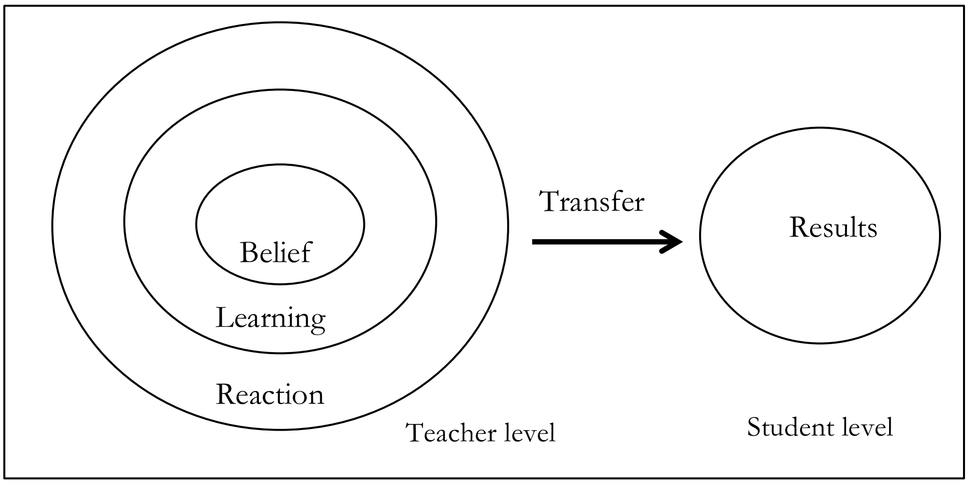

Figure 1: The five level evaluation model

Aim of the study

This research discusses several models about professional development of teachers within the new formats. The aim of this study is evaluating a CPD program that developed for improving teachers' views, beliefs, and practices about NOS, by using the "five level evaluation model" and to introduce this new evaluation perspective.

Method

Design and procedure

This study focuses on a large-scale teacher professional development project aimed at supporting and improving middle school in-service science teachers' professional NOS views, beliefs, and practices. The project began in January 2013 and consisted of a preparation stage and an implementation stage.

During the preparation stage, the CPD's program of NOS workshops, NOS activities, and NOS themes to be emphasised in the activities was planned. Five themes (empirical NOS, tentative NOS, inference and theoretical entities in science, the subjective and theory-laden NOS, and imagination and creativity in science) were derived from: (i) the general thematic structure of the "*Views on Nature of Science Questionnaire, Form C (VNOS-C)* (Lederman, Abd-El-Khalick, Bell & Schwartz, 2002), (ii) the characteristics of the NOS intended to be developed in this project, and (iii) the analytical frameworks used in several studies examining how various groups (e.g., students, teachers, scientists) understood the NOS (e.g., Irez, 2006). These themes were then emphasised in the workshops and activities.

In the second stage of the project, the model of the CPD program about NOS was implemented for a year with the voluntary attendance of 18 middle school science teachers (11 female, 7 male) and their 613 students.

Although there is no single formula for an effective teacher CPD program, there is some consensus about the factors that contribute to a successful professional development experience (Capps, Crawford & Constas, 2012; Luft & Hewson, 2014). In this research, the factors taken into consideration in the development of CPD program about NOS are coherency with other reform initiatives, high quality instruction, active engagement of teachers, enhancement of both content knowledge and pedagogical content knowledge, provision of sufficient time and other resources, provision of sustained support, ensuring collaboration, provision of opportunity for reflection and providing feedback throughout the process, inclusion of evaluation procedures, and provision of local support (blue arrows around the teacher in Figure 2).

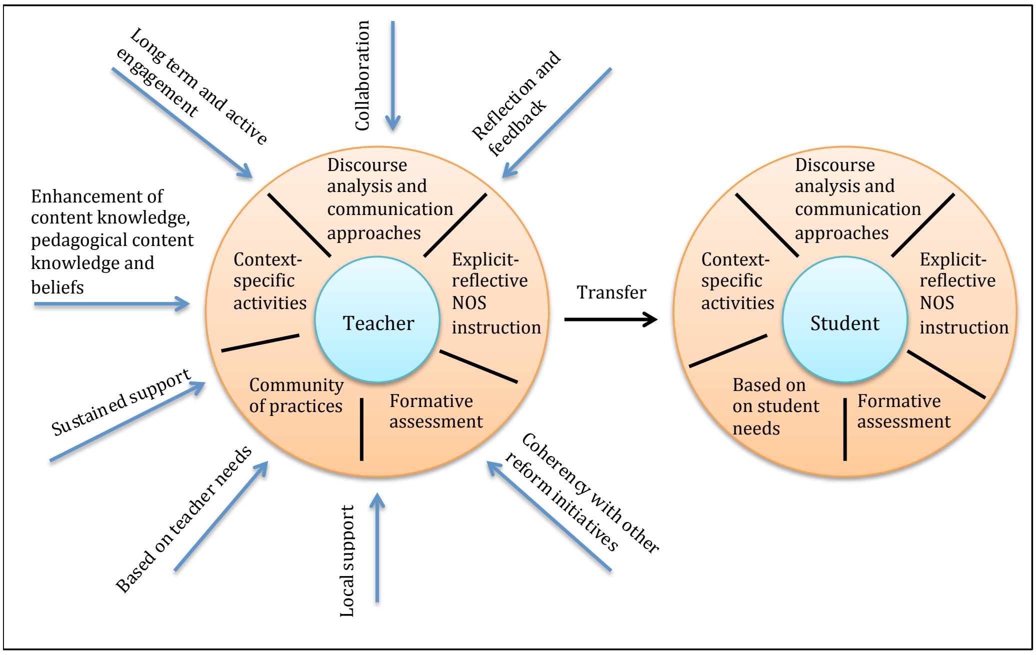

Figure 2: Model of the CPD Program about NOS

As illustrated in Figure 2, the CPD program about NOS included also several teaching approaches. First, explicit-reflective NOS instruction; second, use of *context-specific* activities; third, use of formative assessment; fourth, sustained and long-term engagement with the participating teachers; and fifth, discourse analysis of instruction. The orange circle around the teacher in Figure 2 indicates these teaching approaches taken into account in the periodic workshops organised and the orange circle around the student in Figure 2 indicates these teaching approaches that teachers are expected to adopt in classroom practices. These approaches were used together in a nested manner throughout the professional development program.

The intervention process with the science teachers in this research consisted of 10 monthly meetings (periodic workshops), each lasting about 8 hours, over 2 semesters. Research has demonstrated that explicit-reflective instruction in teaching the NOS is typically more effective than implicit instruction and that NOS instruction is more effective when *context-specific* activities are used rather than *generic* activities (Khishfe & Abd-El-Khalick, 2002). Thus, explicit-reflective instruction was adopted in the teacher training sessions and in all activity materials developed during the project. During the workshops, teachers were introduced to context-specific NOS activities and the teachers' opinions about the activities were recorded. Teachers were given the flexibility to apply the shared activities according to the needs of the students in their classes. The next workshop allowed them to reflect on their experiences and thoughts on their practice. Activities were reorganised according to the views of and suggestions from the teachers. During this process, 57 NOS activities were developed in collaboration with the participant teachers, each met current curriculum guidelines and NOS tenets. The activities are available at Dogan et al. (2016). Research has also shown that using formative assessment rather than summative assessment can improve learning (Bennett, 2011). For this reason, in the CPD process teachers were introduced, in a collaborative and reflective environment, to various NOS aspects and ways of using formative assessment in their NOS teaching. In addition, formative assessments were included in all activity materials provided to the teachers. The teachers were also introduced to different patterns of discourse and communication approaches by analysing video recordings of their classrooms. Sustained and long-term engagement and discourse analyses of instruction were utilised to support teachers in their development in teaching and learning the NOS, and in using explicit instruction and formative assessment. This study is the first to use formative assessment and discourse analysis in a professional development program for in-service teacher education.

Data collection tools and data analysis methods

In this study, data were collected and analysed based on the *"five level evaluation model"*. These levels are learning (teachers' learning about content), beliefs (whether the professional development program makes a difference in teachers' beliefs about teaching and learning content), transfer (whether the professional development program make a difference in teachers' classroom practice), results (student learning outcomes), and reaction (teachers' reactions to the professional development program). Data collection tools and data analysis methods used in the study are discussed below.

Teachers' learning about NOS

Teachers' learning about the NOS was assessed through pre-post semi-structured interviews using "*Views on Nature of Science Questionnaire, Form C (VNOS-C)*" developed by Abd-El-Khalick (1998). Analyses of the interviews were carried out in several steps. First, interviews were transcribed. Second, these transcripts were transferred to the qualitative data analysis program. Third, teachers' statements were grouped regarding NOS themes (the empirical NOS, the tentative NOS, inference and theoretical entities in science, the subjective and theory-laden NOS, and imagination and creativity in science). Before classifying all teacher statements about related themes, inter-rater reliability was also checked. Two participant transcripts were given to two raters who independently classified them. Inter-rater reliability was found to be 82%. Differences were reconciled through discussion between the raters; then, all the teacher statements were classified. At the end of the analysis, teacher statements about related themes were classified as "naive," "eclectic," and "informed." The naive category means having insufficient views on the NOS theme. The eclectic category means having inconsistent and often conflicting views on the concerned NOS theme. The informed category means having consistent views with current approaches to the concerned NOS theme. For aiding the classifying procedure, the rubric developed by Irez (2004) was used, defining each of these categories for each theme (Appendix 1).

Teachers' beliefs about learning and teaching NOS

Teachers' beliefs about teaching and learning NOS were measured by using two Likert-type scales for the pre-post tests, *Self-Efficacy Beliefs Toward Teaching Nature of Science Scale* and the *Science Education and Teaching Belief Scale (BARSTL).*

The Self-Efficacy Beliefs toward Teaching Nature of Science Scale was used for assessing teachers' self-efficacy beliefs as pre-post test. We developed this scale by adapting the *Elementary Science Teaching Efficacy Belief Instrument* developed by Enochs and Rings (1990). To investigate validity and reliability, an initial form was piloted with 328 pre-service science teachers. Exploratory factor analysis was performed to explore the internal structure of the self-efficacy scale. Because of the interval nature of the instrument, polychoric correlations were produced for factor analysis instead of Pearson product moment correlations. Polychoric correlations were used for both determining the number of factors and extracting factors. In order to accomplish that, firstly, the diagonal values of the correlation matrix were replaced by squared multiple correlations in order to approximate the communality estimates. Secondly, parallel analysis was conducted to determine number of factors. The Cronbach alpha coefficient of the full scale was found as .84. (Erdas Kartal, Dogan, Irez, Cakmakci & Yalaki, in-press).

The scale consists of 18 items (11 positive, 7 negative items) distributed under four dimensions: (1) *willingness to teach* NOS (4 items); (2) *personal understanding of* NOS (five items); (3) *pedagogical content knowledge for teaching* NOS (four items); and (4) *assessment of learning* (five items). The responses to the items are recorded on a four-point Likert-type frequency response scale. In scoring, each item response is allocated 1 (Strongly disagree), 2 (Disagree), 3 (Agree), and 4 (Strongly agree) points for each of the response categories. Negative items are scored in reverse and, during analysis, are adjusted accordingly. Possible scores ranged from 18 to 72 points. A higher score indicates higher self-efficacy in teaching the NOS and a lower score represents lower self-efficacy in teaching the NOS. A total test score for each participant teacher was calculated as pre-post test. Pre-test and post-test averages of the teachers were compared at the $p < 0.05$ significance level using the Wilcoxon test. In order to compare the pre-test and post-test averages from the factors of the scale, four Wilcoxon tests were performed, in which case the p value was set to 0.05 / 4 = 0.125 (Bonferonni adjustments) so as not to raise the second type of error rate due to multiplicity.

The Beliefs About Reformed Science Teaching and Learning (BARSTL) Scale developed by Sampson and Benton (2006) was used as pre-post test for assessing teachers' beliefs about reform approaches. The BARSTL Scale includes 16 items that reflect a constructivist science education strategy, and 16 items that reflect a traditional science education strategy. Teachers indicated the degree to which they agreed or disagreed with each of these items using a Likert-type response scale. The items that represent a reformed perspective of science education are scored as 0, 1, 2, 3, respectively, for the responses: Strongly disagree (SD), Disagree (D), Agree (A), and Strongly agree (SA), while the items that represent a traditional perspective are scored in reverse. Possible scores ranged from 0 to 96 points, with higher scores reflecting beliefs about the teaching and learning of science that are more consistent with the current reform movement in science education (as described in AAAS, 1993; National Research Council [NRC], 1996). A total test score for each participant teacher was calculated as pre-post test. At the end, Wilcoxon signed ranks test (2-tailed) was used to compare the results.

Teachers' classroom practices in NOS

Improvement of teachers' practices in the NOS was assessed by video recording during their classroom practices. Thirty-nine video records were obtained and analysed using content analysis. The *Nature of Science Classroom Observation and Artifact Protocol* (NOS-COP) (Herman, Clough & Olson, 2012) was translated into Turkish and was adapted to analyse the data, using ratings "naive", "eclectic" and "informed".

Student outcomes in NOS

Students' (613 students) NOS views were measured by using the *Views on Nature of Science Questionnaire, Form D (VNOS-D)* (Lederman & Khishfe, 2002) as post-tests. In data analysis, five themes of the NOS (empirical NOS, tentative NOS, inference and theoretical entities in science, subjective and theory-laden NOS, imagination and creativity in science) were chosen to assess changes in the views of students. The data was classified as "naive-1", "eclectic-2", or "informed-3". To assess the student responses and to make the relevant coding, the researchers relied on the rubric developed by Lederman and Holliday (2011). Wilcoxon signed ranks test (2-tailed) was used to compare the results.

Teachers' reactions to the NOS-CPD program

Teachers' reactions to the CPD program were collected through interviews at the end of the study by using five open-ended questions. The first question concerned contributions of the professional development programs to teaching as a profession. The second question inquired whether the professional development programs had changed teachers' classroom practices or their perspectives upon teaching. The third question asked about differences between the professional development program and other programs they may have attended. The fourth question sought to identify strengths of the professional development program. The last question sought information on aspects that were lacking or could be improved. Data were analysed by using content analysis in six dimensions (the program's impact on teachers' knowledge about the NOS; impact on teachers' beliefs about teaching and learning the NOS; impact on teachers' classroom practices; program comparison; strengths of the program; and weaknesses of the program).

Results

Teachers' views about NOS

Research findings showed that the CPD program about NOS improved teachers' views toward the NOS. As seen in Table 1, the ratio of teachers who had naive views about the NOS themes of the program decreased over the course of the training, whereas the ratio of teachers sharing informed views increased (Appendix 2).

As evident in the table, teachers made more progress in *"imagination and creativity in science"* theme than other themes in the table. Most of the teachers thought that imagination and creativity were indispensable in the production of scientific knowledge before the professional development program; however, they stated that, instead of being used at every stage of the process, they were used during the beginning and the process, in the conclusion part the scientists should be away from creativity and imagination and be objective. However, after the professional development program, most of the teachers adopted that imagination and creativity were effective at and used in every stage of the production of scientific knowledge:

> The result is affected by the experiments. Imagination and creativity were initially effective when establishing hypotheses. (Duru_pre interview).

> And my views on this have changed, obviously. I think at the moment (imagination and creativity) can use it at every stage. (Duru_post interview)

On the other hand, in "subjective and theory-laden NOS" teachers showed less improvement than predicted, compared to other themes. While most of the teachers had naive opinions about this theme before the professional development program, they passed the eclectic view after the professional development program. Prior to the professional development program, these teachers stated that indirect observations and modeling did not have a place in science. Most of these teachers stated that at the end of the program, direct observation can be used in the production of scientific information as well as indirect observation, and modeling can be used in science. But their views are not informed. Because the teachers who could not reach the stated opinion in this theme could not realise that theories could be based upon indirect observations and predictions. They are also not aware that models are created by people to facilitate understanding of events:

> They have seen the structure of the atom, because we have solved the structure of the atom. If they have not seen the structure we already call them in our lessons. I'm saying they've done some important work there. (Buse_pre interview)

> Of course, we cannot observe atoms; but we can reveal their existence from natural events. This is the most beautiful experiment in the classroom, we rub the balloon to the ceiling, and we're sticking to the ceiling. It is not direct in this way; but even if it is indirect, we say that we can observe this way. (Buse_post interview)

Teachers' beliefs about learning and teaching NOS

Research findings showed that the CPD program about NOS positively improved teachers' self-efficacy beliefs toward teaching the NOS (Table 2). There was a significant difference between teachers' total pre-test and post-test scores ($p<0.05$) (Appendix 3).

As seen in Table 3, according to the Wilcoxon signed ranks test results for each sub-dimension of *The Self-Efficacy Beliefs Toward Teaching Nature of Science Scale*, the difference between the pre-test and post-test scores were not significant at $p > 0.0125$ (Appendix 4). Possible reasons for this are presented in the discussion.

The CPD program about NOS positively improved teachers' beliefs toward reform approaches, as there was a significant difference between teachers' total pre-test and post-test scores ($p<0.05$) (Appendix 5: Table 4).

Teachers' classroom practices in NOS

The CPD program about NOS improved teachers' classroom practices of the NOS (Tekin, Dogan, Irez, Yalaki & Cakmakci, 2019). Table 5 shows the naive, eclectic, or informed rates of teachers' expressions about the NOS within the classroom (Appendix 6). The findings indicate that although the percentages in the "naive" and "informed" categories fluctuated from month to month throughout the process, it is evident that the ratio in the "naive" category decreased and the ratio in the "informed" category increased at the end of the program (Table 5).

Student outcomes in NOS

The CPD program about NOS improved students' views about the NOS (Ozer, Doğan, Yalaki, Irez & Yalaki, 2019). The Wilcoxon signed rank test was carried out between the students' pre-post test performance to determine whether their development in the NOS themes was statistically significant (Table 6). There was a significant difference between the pre-post test scores of 5th, 6th and 7th grade students in the empirical NOS, tentative NOS, inference and theoretical entities in science, and subjective and theory-laden NOS themes. A meaningful difference was found between the pre-post test performance of 7th grade students in terms of imagination and creativity in science theme. There was no significant difference between pre-post test performances of 5th and 6th grade students regarding this theme.

The CPD program also positively improved students' NOS views in all grades throughout the process (Table 7, Appendix 8). The percentages of students who had naive views at each grade level decreased, and the percentages of students who had eclectic and informed NOS views increased. Fifth grade students showed the most improvement in the inference and theoretical entities in science theme, whilst NOS theme that 5th grade students developed least was creativity and imagination in science. Sixth grade students showed the most improvement in the tentative NOS theme, but were the least developed in imagination and creativity in science. Seventh grade students showed the most improvement in the themes empirical NOS, and imagination and creativity in science.

Teachers' reactions to the NOS-CPD Program

As seen in Table 8, teachers generally thought the CPD program about NOS was effective in improving their professional competences about the NOS (Appendix 9). Half of the teachers said that the professional development program improved their content knowledge about NOS:

> I have learnt a lot of things in this project. I have noticed that I did not have much information regarding the NOS and learnt about it thanks to this program. (Harun)

Similarly, more than half of the teachers indicated that the program helped to improve their classroom practices about NOS. For example, 61% of the teachers shared that they made a progress about integrating the NOS into scientific content:

> We give more to our students than just teaching them the subject. In other words, we give information about how science works. We don't just teach. Thus, I think the program has made a difference in this regard. (Lara)

During the research, teachers were asked to compare the current CPD program with other CPD programs in which they had participated. In their answers, the teachers mostly mentioned *the social interaction opportunities* in the CPD program about NOS. In this regard, 33% of the participants said that these types of opportunities make it much better than other programs:

> We had an opportunity to come together with other science teachers which is one of the strengths of this program. I think the communication we established was very good. (Gamze)

Teachers were also asked to evaluate the strengths and weaknesses of the CPD program about NOS. With regards to the strengths of the program, the teachers mostly mentioned the exchange of information and experiences among teachers, the consideration of the feedback received from teachers in the activities, the provision of feedback about teacher practices, and opportunity to participate actively. In this regard, 56% of teachers indicated that *the exchange of information and experience among teachers* was one of the strengths that came to the forefront.

> There, we exchanged information with science teachers. For instance we talked about which unit we were currently teaching, how we were applying some methods, etc. (Buse)

When the weaknesses of the CPD program about NOS were evaluated, the participants generally avoided mentioning any. Only a few teachers (17%) pointed to the fact that *the language used in some activities was above the level of students*. One teacher indicated the following weaknesses of the program: the workshops took a whole day and they were very few in number; scientific gains stayed in the background in some activities; there were no step taken to increase participation; the video shooting had a negative effect on the classroom participation of students; and the teachers were not given an opportunity to write the activities.

> It would be better if you had given us a subject and asked us to prepare an activity including the NOS. If we had been asked to prepare an activity and make a presentation of it, we would observe better the changes in ourselves. (Akın)

Discussion

Professional development programs for teachers have become an important component of educational policies for increasing the quality of teaching and learning in schools. Funding for these professional development programs is regulated by state budgets and also can come from various other sources. As investments in professional development have increased, politicians have begun to ask for evidence of how professional development programs are affecting learning outcomes, not just about the impact of professional development on teacher knowledge and practice. This indicates that there is a need for more sophisticated models of assessment that will support the assessment of professional development programs at multiple levels (Ingvarson, Meiers & Beavis, 2005). However, a literature review shows that while many different types of CPD programs have been implemented recently and various evaluation models suggested, researchers have not been paying enough attention to the evaluation of their programs (Guskey, 2000). More emphasis should be laid on the evaluation of professional development programs, which will address a significant gap in the literature.

Starting from this gap in the literature, effectiveness of the CPD program about NOS was evaluated in this study by using a sophisticated model that allows for the assessment of professional development programs at multiple levels. The findings obtained from this study demonstrate that the CPD program about NOS effectively improved teachers' views about the NOS. Before the CPD program about NOS, teachers had naive and eclectic views in almost all the NOS themes. After the CPD program about NOS, there was a significant decrease in the percentage of teachers with naive and eclectic views and there was a significant increase in the percentage of teachers with informed views in the almost all the relevant NOS themes (Table 1). Moreover, the findings obtained from this study demonstrate that the CPD program about NOS effectively improved teachers' beliefs about learning and teaching the NOS.

One finding is particularly noteworthy. There was a significant difference between teachers' total pre-post test scores on *The Self-Efficacy Beliefs Toward Teaching Nature of Science Scale* at $p<0.05$, but the difference between the pre-test and post-test performances of teachers obtained from the scale's dimensions were not significant at $p <0.0125$ (Table 2 and Table 3). This may be due to the fact that the pre-test averages of teachers are higher than expected. As they were not aware of misconceptions and inadequate concepts that they held before the CPD program about NOS, their self-efficacy toward teaching the NOS remained quite high. Teachers who lack essential knowledge and understanding of the NOS but who have high self-efficacy and confidence toward teaching NOS could be the major barriers in promoting scientific literacy. Their willingness to confidently teach inadequate NOS concepts will inevitably result in their students leaving formal education without an informed understanding of the NOS. For this reason, it is important that teachers are confronted with their own misconceptions which professional development programs should eliminate. The findings obtained from this study demonstrated that the CPD program about NOS enhanced teachers' classroom practices about NOS. Similarly, findings showed that the CPD program about NOS improved students' views about the NOS. Teachers' reaction about the effectiveness of the CPD program about NOS also supports these findings.

The literature supports the results that participation in long-term professional development programs improves teachers' views about NOS (Akerson & Hanuscin, 2007), beliefs about learning and teaching the NOS (Bell & Maeng, 2013) and classroom practices about the NOS (Akerson & Hanuscin, 2007), and these changes reflect positively on learning outcomes for students (Lederman et al., 2012). When the findings obtained from the research were evaluated separately, the expected results were obtained in accordance with the literature. When the research findings were evaluated together, students' views about NOS themes improved less than teachers' views about NOS themes (Table 1 and Table 7) did. This may be due to the fact that teachers' classroom practices are less developed than their NOS views (Table 1 and Table 5). Effectiveness of the CPD program about NOS is related to teacher profiles and classroom practices that improve the views of students. For this reason, Ozer and her colleagues (2018) argued that teachers should focus more on monitoring classroom practices throughout their professional development program. On the other hand, teachers' classroom practices are not easily predictable. The gains that teachers make in professional development programs are influenced by other variables (competencies, beliefs, identities, and missions) (Korthagen, 2004). More research needs to be done that takes into account the variables that affect classroom practices of teachers to increase the effectiveness of professional development programs.

Last of all, as mentioned above, evaluation of the effectiveness of professional development programs is very complex and problematic. This situation may cause researchers to be reluctant to pay appropriate attention to the evaluation of their programs. This work presents a new perspective for evaluating professional development programs, which may guide researchers to multidimensional evaluation of their programs. In this study we used the five level evaluation model to evaluate the effectiveness of the CPD program about NOS, but we think that this model can be used to evaluate the professional development programs in any subject. It is important, however, to note that the utility of "five levels" is very dependent upon good preparation of new instruments and also on the good use of piloting or a good selection from existing instruments.

Acknowledgements

This paper is part of the professional development project supported by TUBITAK under 111K527 project number. This paper is also a part of first author's PhD thesis and a part of this study was presented at the National Association for Research in Science Teaching (NARST) (2016), Baltimore, USA.
https://narst.org/annualconference/NARST_2016_Abstracts.pdf

References

Abd-El-Khalick, F., Bell, R. L. & Lederman, N. G. (1998). The nature of science and instructional practice: Making the unnatural natural. *Science Education,* 82, 417-436. https://doi.org/10.1002/(SICI)1098-237X(199807)82:4%3C417::AID-SCE1%3E3.0.CO;2-E

Akerson, V. L. & Hanuscin, D. L. (2007). Teaching nature of science through inquiry: Results of a 3-year professional development program. *Journal of Research in Science Teaching*, 44(5), 653-680. https://doi.org/10.1002/tea.20159

Bell, R. L. & Maeng, J. L. (2013). Statewide professional development to support reforms-based science instruction: Results from two years of implementation. Paper presented at the *Annual Meeting of the Association for Science Teacher Education*, Charleston, SC. https://vista.gmu.edu/assets/docs/research/NARST_VISTA_2013.pdf

Bennett, R. E. (2011). Formative assessment: A critical review. *Assessment in Education: Principles, Policy & Practice,* 18(1), 5-25. https://doi.org/10.1080/0969594X.2010.513678

Burton, E. P. (2013). Student work products as a teaching tool for nature of science pedagogical knowledge: A professional development project with in-service secondary science teachers. *Teaching and Teacher Education,* 29, 156-166. https://doi.org/10.1016/j.tate.2012.09.005

Capps, D. K., Crawford, B. A. & Constas, M. A. (2012). A review of empirical literature on inquiry professional development: Alignment with best practices and a critique of the findings. *Journal of Science Teacher Education,* 23(3), 291-318. https://doi.org/10.1007/s10972-012-9275-2

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. https://doi.org/10.1080/09500693.2012.760209

Cheng, M. M. H., Chan, K. W., Tang, S. Y. F. & Cheng, A. Y. N. (2009). Pre-service teacher education students' epistemological beliefs and their conceptions of teaching. *Teaching and Teacher Education*, 25(2), 319-327.

Dass, P. M. & Yager, R. E. (2009). Professional development of science teachers: History of reform and contributions of the STS-based Iowa Chautauqua Program. *Science Education Review,* 8(3), 99-111. https://www.researchgate.net/publication/228492636_Professional_Development_of_Science_Teachers_History_of_Reform_and_Contributions_of_the_STS-Based_Iowa_Chautauqua_Program

Dogan, N., Cakmakci, G., Irez, S., Yalaki, Y., Erdas, E., Kaya, G., Ozer, F., Han, Tosunoglu, C., Altiner, Z. B., Bala, V. G., Ertugrul, G., Sardag, M., Koylu, N. Z. & Gunsever, G. (2016). In Y. Yalaki (Ed.), *Etkinliklerle Bilimin Doğasının Öğretimi* (Genisļetilmiş 2. Baskı). Pegem Akademi Yayıncılık, Ankara. http://www.pegem.net/dosyalar/dokuman/191701_ekitap.pdf

Enochs, L. G. & Riggs, I. M. (1990). Further development of an elementary science teaching efficacy belief instrument: A preservice elementary scale. *School Science and Mathematics,* 90(8), 694-706. https://doi.org/10.1111/j.1949-8594.1990.tb12048.x

Erdaş, E. (2015). *Supporting the teachers' professional developments though the process about teaching NOS: A continuing professional development program model.* Unpublished PhD Thesis, Abant Izzet Baysal University, Turkey.

Erdas Kartal, E., Cobern, W. W., Dogan, N., Irez, S., Cakmakci, G. & Yalaki, Y. (2018). Improving science teachers' nature of science views through an innovative continuing professional development program. *International Journal of STEM Education,* 5:30. https://doi.org/10.1186/s40594-018-0125-4

Erdas Kartal, E., Dogan, N., Irez, S., Cakmakci, G. & Yalaki, Y. (in-press). Science teachers' beliefs about learning and teaching nature of science: Professional development program**.** *Education and Science.*

Fenstermacher, G. D. & Berliner, D. C. (1983). *A conceptual framework for the analysis of staff development.* Unpublished manuscript, Santa Monica, California. https://www.rand.org/content/dam/rand/pubs/notes/2007/N2046.pdf

Fullan, M. G. (1995). The limits and the potential of professional development. In T. Guskey & M. Huberman (Eds.), *Professional development in education: New paradigms and practices*, 253-267.

Guskey, T. R. (2000). *Evaluating professional development.* Thousand Oaks, CA: Corwin Press.

Guskey, T. R. (2003). What makes professional development effective? *Phi Delta Kappan* 84(10), 748-750. https://doi.org/10.1177%2F003172170308401007

Guskey, T. R. & Yoon, K. S. (2009). What works in professional development? *Phi Delta Kappan,* 90(7), 495-500. https://doi.org/10.1177%2F003172170909000709

Haney, J. J. & Lumpe, A. T. (1995). A teacher professional development framework guided by reform policies, teachers' needs, and research. *Journal of Science Teacher Education*, 6(4), 187-196. https://doi.org/10.1007/BF02614642

Herman, B. C., Clough, M. P. & Olson, J. K. (2013). Teachers' nature of science implementation practices 2–5 years after having completed an intensive science education program. *Science Education,* 97(2), 271-309. https://doi.org/10.1002/sce.21048

Ingvarson, L., Meiers, M. & Beavis, A. (2005). Factors affecting the impact of professional development programs on teachers' knowledge, practice, student outcomes & efficacy. *Education Policy Analysis Archives,* 13(10), 1-26. https://epaa.asu.edu/ojs/article/view/115/241

Irez, O. S. (2004). *Turkish pre-service science teacher educators' beliefs about the nature of science and conceptualizations of science education*. Unpublished PhD dissertation, Nottingham University, UK.

Irez, S. (2006). Are we prepared?: An assessment of preservice science teacher educators' beliefs about nature of science. *Science Education*, 90(6), 1113-1143. https://doi.org/10.1002/sce.20156

Kanadlı, S. (2012). *Investigating the effectiveness of a professional development program designed for teachers.* PhD Thesis, Gazıantep University, Turkey.
Kelchtermans, G. (2004). CPD for professional renewal: moving beyond knowledge for practice. In C. Day & J. Sachs (Eds.), *International handbook on the continuing professional development of teachers* (pp. 217-237). Berkshire: Open University Press.
Khishfe, R. & Abd-El-Khalick, F. (2002). Influence of explicit and reflective versus implicit inquiry-oriented instruction on sixth graders' views of nature of scientific knowledge. *Journal of Research in Science Teaching,* 39(7), 551-578. https://doi.org/10.1002/tea.10036
Kirkpatrick, D. L. (1959). Techniques for evaluating training programs. *Training and Development Journal,* 3(11), 1-13.
Kirkpatrick, D. (2001). The four-level evaluation process. In L. L. Ukens (Ed.), *What smart trainers know: The secrets of success from the world's foremost experts* (pp. 122-132). San Francisco, CA: Jossey-Bass/Pfeiffer.
Kirkpatrick, D. L. & Kirkpatrick, J. D. (2007). *Implementing the four levels.* Berret-Koehler Publishers (p. 47), San Francisco.
Korthagen, F. A. J. (2004). In search of the essence of a good teacher: Towards a more holistic approach in teacher education. *Teaching and Teacher Education,* 20(1), 77-97. https://doi.org/10.1016/j.tate.2003.10.002
Koseoglu, F., Tumay, H. & Ustun, U. (2010). Developing a professional development package for nature of science instruction and discussion about its implementation for pre-service teachers. *Journal of Kirsehir Education Faculty*, 11(4)*,* 129-162. http://journaldatabase.info/articles/developing_professional_development.html
Lederman, N. G., Abd-El-Khalick, F., Bell, R. L. & Schwartz, R. S. (2002). Views of the nature of science questionnaire: Toward valid and meaningful assessment of learner's conceptions of the nature of science. *Journal of Research in Science Teaching,* 39(6), 497-521. https://doi.org/10.1002/tea.10034
Lederman, J. S. & Khisfe, R. (2002). Views of the nature of science, Form D. Unpublished paper. Illinois Institute of Technology, Chicago, IL.
Lederman, J. S., Lederman, N. G., Kim, B. S. & Ko, E. K. (2012). Teaching and learning of nature of science and scientific inquiry: Building capacity through systematic research-based professional development. In M. S. Khine (Ed.), *Advances in the nature of science research: Concepts and methodologies* (pp.125-151). Dordrecht: Springer.
Lowden, C. (2005). Evaluating the impact of professional development [Electronic Version]. *The Journal of Research in Professional Learning.*
Luft, J. A. & Hewson, P. W. (2014). Research on teacher professional development programs in science. In N. Lederman & S. Abell (Eds.), *Handbook of research on science education* (pp. 889-909). Taylor and Francis.
Lumpe, A., Czerniak, C., Haney, J. & Beltyukova, S. (2012). Beliefs about teaching science: The relationship between elementary teachers' participation in professional development and student achievement. *International Journal of Science Education,* 34(2), 153-166. https://doi.org/10.1080/09500693.2010.551222
Mesci, G. & Schwartz, R. S. (2017). Changing pre-service science teachers' views of nature of science: Why some conceptions may be more easily altered than others? *Research in Science Education,* 47(2), 329-351. http://dx.doi.org/10.1007/s11165-015-9503-9

Metfessel, N. S. & Michael, W. B. (1967). A paradigm involving multiple criterion measures for the evaluation of the effectiveness of school programs. *Educational and Psychological Measurement*, 27(4), 931-943. https://doi.org/10.1177%2F001316446702700441

Mitchener, C. P. & Anderson, R. D. (1989). Teachers' perspective: Developing and implementing an STS curriculum. *Journal of Research in Science Teaching*, 26(4), 351-369. https://doi.org/10.1002/tea.3660260407

Nespor, J. (1987). The role of beliefs in the practice of teaching. *Journal of Curriculum Studies*, 19(4), 317-328. https://doi.org/10.1080/0022027870190403

Özer, F., Doğan, N., Yalaki, Y., İrez, S. & Çakmakcı, G. (2019). The ultimate beneficiaries of continuing professional development programs: Middle school students' nature of science views. *Research in Science Education,* 1-26. https://doi.org/10.1007/s11165-019-9824-1

Posnanski, T. J. (2002). Professional development programs for elementary science teachers: An analysis of teacher self-efficacy beliefs and a professional development model. *Journal of Science Teacher Education*, 13(3), 189-220. https://doi.org/10.1023/A:1016517100186

Renyi, J. (1996). Teachers take charge of their learning. Transforming professional development for student success [and] executive summary. National Foundation for the Improvement of Education, Washington, DC. https://eric.ed.gov/?id=ED401251

Sampson, V. & Benton, A. (2006). Development and validation of the beliefs about science teaching and learning (BARSTL) questionnaire. Paper presented at the annual meeting of the Association for Science Teacher Education, Portland, Oregon, USA.

Scriven, M. (1967). The methodology of evaluation. In R. E. Stake (Ed.), *Curriculum evaluation.* American Educational Research Association Monograph Series on Evaluation, No. I - Chicago: Rand McNally.

Stufflebeam, D. L. (1969). Evaluation as enlightenment for decision-making. In A. Walcott (Ed.), *Improving educational assessment and an inventory of measures of affective behavior.* Washington, DC: Association for Supervision and Curriculum Development.

Tekin, G., Doğan, N., İrez, S., Yalaki, Y. & Çakmakcı, G. (2019). The impact of a professional development program on science teachers' classroom performance: Teaching nature of science. *PAU Journal of Education,* 46(41).

Tyler, R. W. (1942). General statement on evaluation. *The Journal of Educational Research,* 35(4), 492-501. https://doi.org/10.1080/00220671.1942.10881106

Worthen, B. R. & Sanders, J. R. (1987). *Educational evaluation: Alternative approaches and practical guidelines.* Longman: New York.

Appendix 1 : Rubric for coding teachers' NOS views (Irez, 2004)

Themes	Naive	Eclectic	Informed
Empirical NOS	Describes science as being solely dependant on direct evidence, believes that scientific claims can (only) be proven by direct evidence.	Believes that science solely relies on direct evidence but accepts that evidence supports rather than proves scientific claims. Believes that science does not solely rely on direct evidence but accepts that evidence proves scientific claims.	Believes science uses both direct and indirect evidence and claims that evidence supports rather than proves scientific claims.
Scientific method	Believes that there is a single universal scientific method which scientists follow step-by-step to reach conclusions.	Believes that there exists a universal scientific method which is not a stepwise procedure.	Believes that there are many methods in science and sees method as related to paradigm.
Tentative NOS	Claims that scientific knowledge is true and certain.	Accepts that some scientific theories are tentative but claims that scientific laws are true and not subject to change.	Believes that all scientific knowledge, regardless to its nature or status, is subject to change and modifications in the future.
Nature of scientific theories and laws	Believes that theories are not well sustained and therefore subject to change. Also claims that, when proven, theories become laws which have higher status and are not subject to change.	Believes in the well-sustained nature of theories, thus views them as subject to change. However, fails to recognise theories and laws as different kinds of scientific knowledge or believes that laws have higher status and are not subject to change.	Believes that theories are well-supported explanation systems. Demonstrates an understanding that theories and laws are different kinds of scientific knowledge, and laws, as well as theories, are subject to change.
Inference and theoretical entities in science	Believes in science's reliance on direct evidence and therefore does not appreciate the inferential nature of some theories.	Although accepting reliability of some theories which are based on inference, objects to some others claiming that there is no direct evidence to support (or prove) them.	Demonstrates a comprehensive understanding of the inferential nature of some theories.

Themes	Naive	Eclectic	Informed
Subjective and theory-laden NOS	Believes that science pictures an objective account of nature due to its methods and objectivity of its practitioners.	Believes that there could be differences amongst scientists in data interpretation due to their professional backgrounds. Believes that there could be differences amongst scientists in data interpretation due to their personal values and beliefs.	Views subjectivity as integral to the construction of scientific knowledge and believes that scientists' professional and personal backgrounds cause subjectivity.
Social and cultural embedd-edness of science	Claims that science is universal and denies social and cultural influences on science.	Accepts that society and culture affect some scientific disciplines (such as evolutionary biology), but not all (such as chemistry).	Believes that science affects and is affected by society and culture.
Imagin-ation and creativity in science	Rejects that science involves imagination and creativity.	Believes that certain stages of scientific inquiry involve imagination and creativity. However, holds inconsistent views about scientific methodology and the inferential nature of some scientific theories.	Believes that imagination and creativity permeates the scientific process throughout.

Appendix 2: Table 1: Percentage of teachers' NOS views in the pre-post interview

	Naive		Eclectic		Informed	
	Pre (%)	Post (%)	Pre (%)	Post (%)	Pre (%)	Post (%)
Empirical NOS	55	11	39	34	6	55
Tentative NOS	6	0	78	39	16	61
Inference and theoretical entities in science	44	6	34	55	22	39
Subjective and theory-laden NOS	22	0	50	39	28	61
Imagination and creativity in science	0	0	61	11	39	89

N = 18

Appendix 3: Table 2: Wilcoxon signed ranks test results for self-efficacy beliefs toward teaching NOS Scale

	Post > Pre	Post = Pre	Post < Pre	p
Pre-test*Post test	11	2	5	.021*

* $p < 0.05$; n = 18
PostT > PreT: Number of teachers whose post-test score is higher than the pre-test score
PostT = PreT: Number of teachers whose post-test score equals the pre-test score
PostT < PreT: Number of teachers whose post-test score is lower than the pre-test score

Appendix 4: Table 3: Wilcoxon signed ranks test results for each sub-dimension of the self-efficacy beliefs toward teaching NOS Scale

Dimensions	Pre-test mean	Post-test mean	PostT > PreT	PostT = PreT	PostT < PreT	p
Personal self-efficacy belief about teaching the NOS	11.6	12.6	9	7	2	.015
Teacher role on the teaching the NOS	16.5	16.6	9	0	9	.982
Self-efficacy belief about teaching process	12.1	12.7	8	6	4	.080
Self-efficacy belief about assessment process	14.6	16.1	8	6	4	.044

* $p < 0.0125$; N = 18
PostT > PreT: Number of teachers whose post-test score is higher than the pre-test score
PostT = PreT: Number of teachers whose post-test score equals the pre-test score
PostT < PreT: Number of teachers whose post-test score is lower than the pre-test score

Appendix 5: Table 4: Wilcoxon signed ranks test results for the BARSTL Scale

	Post > Pre	Post = Pre	Post < Pre	p
Pre-test*Post test	15	1	2	.001*

* $p < 0.05$; N = 18
PostT > PreT: Number of teachers whose post-test score is higher than the pre-test score
PostT = PreT: Number of teachers whose post-test score equals the pre-test score
PostT < PreT: Number of teachers whose post-test score is lower than the pre-test score

Appendix 6: Table 5: Teachers' level of reflecting NOS themes

	Feb %	Mar %	Apr %	May %	Oct %	Nov %	Dec %
Naive	30.8	22.3	16.7	0	50	0	20
Eclectic	53.8	19.4	27.8	20	0	50	20
Informed	15.4	58.3	55.5	80	50	50	80

N = 18

Appendix 7: Table 6: Wilcoxon signed ranks test results for VNOS-D Questionnaire

Themes	5. Grade Pre*Post p	6. Grade Pre*Post p	7. Grade Pre*Post p
Empirical NOS	.000*	.003*	.000*
Tentative NOS	.000*	.000*	.000*
Inference and theoretical entities in science	.000*	.012*	.000*
Subjective and theory-laden NOS	.007*	.002*	.000*
Imagination and creativity in science	.200	.364	.000*

*$p < 0.05$

Appendix 8: Table 7: Change of students' views about NOS themes by grade

Theme	Category	5. Grade		6. Grade		7. Grade	
		Pre %	Post %	Pre %	Post %	Pre %	Post %
Empirical NOS	Naive	60.7	42.3	58.9	46.0	68.3	25.8
	Eclectic	32.5	54.6	38.1	53.1	28.7	64.5
	Informed	0.0	2.1	0.4	0.9	1.2	9.0
Tentative NOS	Naive	47.9	30.9	40.8	21.1	47.0	20.0
	Eclectic	32.5	54.6	48.7	72.3	40.2	65.2
	Informed	0.0	2.1	0.8	1.9	0.6	10.3
Inference and theoretical entities in science	Naive	62.4	43.3	55.5	46.5	52.4	27.1
	Eclectic	22.2	50.5	39.6	48.8	40.2	66.5
	Informed	0.0	2.1	0.4	2.3	0.6	5.2
Subjective and theory-laden NOS	Naive	42.7	33	52.1	39.4	54.3	34.8
	Eclectic	23.1	39.2	37.7	47.4	29.3	49.7
	Informed	0.0	3.1	0.8	5.6	0.0	6.5
Imagination and creativity in science	Naive	23.1	15.5	9.4	9.9	11.0	10.3
	Eclectic	65.8	76.3	73.6	77.5	72.6	73.5
	Informed	0.0	2.1	3.0	8.4	1.2	15.5

N = 613

Appendix 9: Table 8. Teachers' reactions about effectiveness, strengths and weaknesses of the CPD program

Dimensions of evaluation for professional development program	Sub-dimensions of evaluation for professional development program	% (n=18)
Effect of the program on the information fields related to NOS	Content knowledge	50
	Pedagogical content knowledge	78
	Science literacy	17
Effect of the program on beliefs about teaching and learning NOS	Self sufficiency	22
	Curiosity and interest	17
Effect of the program on classroom practices	Laboratory practices	11
	Integrating the NOS into scientific content	61
	Encouraging the student	50
Comparison of the program	Sincere environment	17
	Productive process	17
	Social interaction	33
Strengths of the program	Quality of the activities	11
	Feedback	44
	The exchange of information and experience among teachers	56
	Active participation	44
	Long-term program	28
	Professional support	17

	Taking into consideration the feedback received from teachers, in the activities organised	44
	Fitness for the purpose	6
	Professional progress (disciplined study period)	6
	Activities in parallel with the curriculum	6
	Technological support	11
	Material and resource support	33
	Increasing the motivation of the student	33
	Content-based activities	6
Weaknesses of the program	The duration of the workshops was long	6
	The number of workshops was few	6
	Suitability of the activities to the level of the student	17
	The scientific gains of the activities stayed in the background	6
	Teachers were not asked to write the activities	6
	The actions to increase the participation	6
	The effect of video-shooting on the classroom participation of students	6

Eda Erdas Kartal is an Assistant Professor of Educational Science at Kastamonu University, Turkey. Her research interests include teacher education, teaching and learning nature of science, teaching evolution, and professional development of science teachers. Her research interests also include educational programs and teaching.
Email: erdaseda@gmail.com

Nihal Dogan is an Associate Professor of Science Education at Abant Izzet Baysal University, Turkey. Her research focuses on the teaching and learning of nature of science in grades K-12, and in pre-service and in-service science teacher. Her research interests also include teacher education, professional development of science teachers and research methods in science education.
Email: nihaldogan17@gmail.com

Serhat Irez is a Professor of Science Education at Marmara University, Turkey. His research interests include teaching and learning nature of science, the history of science in science teaching, teaching evolution and, professional development of science teachers.
Email: sirez@marmara.edu.tr

Gultekin Cakmakci is a Professor of Science Education at Hacettepe University, Turkey and has been teaching courses on STEM education and public engagement with STEM. His research interests focus on developing scientific literacy among students and the general public and on the design, implementation and evaluation of STEM teaching.
Email: gultekincakmakci@gmail.com
Web: http://gultekincakmakci.weebly.com

Yalcin Yalaki graduated from Cukurova University, Department of Chemistry. He completed his MS and PhD degrees at Florida State University in science education. He has been working at Hacettepe University since 2004. His research interests include, teaching and learning of nature of science, inquiry based science education, and formative assessment.
Email: yyalaki@gmail.com

Please cite as: Erdas Kartal, E., Dogan, N., Irez, S., Cakmakci, G., Yalaki, Y. (2019). A five-level design for evaluating professional development programs: Teaching and learning about nature of science. *Issues in Educational Research*, 29(2), 402-426. http://www.iier.org.au/iier29/erdas-kartal.pdf