

Central Lancashire Online Knowledge (CLoK)

Title	Teachers' ability to construct arguments, but not their perceived self-efficacy of teaching, predicts their ability to evaluate arguments
Type	Article
URL	https://clock.uclan.ac.uk/31485/
DOI	https://doi.org/10.1080/09500693.2020.1722864
Date	2020
Citation	Lytzerinou, Evangelia and Iordanou, Kalypso (2020) Teachers' ability to construct arguments, but not their perceived self-efficacy of teaching, predicts their ability to evaluate arguments. <i>International Journal of Science Education</i> , 42 (4). pp. 617-634. ISSN 0950-0693
Creators	Lytzerinou, Evangelia and Iordanou, Kalypso

It is advisable to refer to the publisher's version if you intend to cite from the work.
<https://doi.org/10.1080/09500693.2020.1722864>

For information about Research at UCLan please go to <http://www.uclan.ac.uk/research/>

All outputs in CLoK are protected by Intellectual Property Rights law, including Copyright law. Copyright, IPR and Moral Rights for the works on this site are retained by the individual authors and/or other copyright owners. Terms and conditions for use of this material are defined in the <http://clock.uclan.ac.uk/policies/>

Teachers' Ability to Construct Arguments, but not Their Perceived Self-Efficacy of Teaching, Predict Their

Ability to Evaluate Arguments

Evangelia Lytzerinou^a & Kalypso Iordanou^b

^aOpen University of Cyprus
PO Box 12794,
2252, Latsia
Email: elytzerinoy@gmail.com

^bUniversity of Central Lancashire, Cyprus
12-14 University Ave,
Pyla, Larnaca
Cyprus
Email: KIordanou@uclan.ac.uk

Abstract

The aim of the present study was to examine science and non-science education secondary school teachers' skill to evaluate arguments, and how this skill relates to their skill to construct arguments and to their perceptions about their ability to teach argumentation skills effectively. The study also examined whether teachers' arguments skills and their self-efficacy of teaching argumentation were domain specific. Social-science education teachers, who teach literature and history, and physical science education teachers, were asked to write two essays – one on a social topic and another on a socio-scientific topic–, to evaluate the quality of written arguments and to complete an instrument assessing their self-efficacy of teaching argumentation. Results showed that teachers' ability to construct arguments predicted their ability to evaluate arguments. Yet, although teachers expressed high self-efficacy in teaching argumentation in both domains, their abilities to evaluate and construct arguments were not sufficiently developed, in neither domain. The findings of the present study have important educational implications, suggesting that specific attention needs to be paid on teachers' skills of constructing and evaluating arguments.

Keywords: Argument skills; argument evaluation; domain; teachers' perceptions

Teachers' ability to construct arguments, but not their perceived self-efficacy of teaching, predict their ability to evaluate arguments

1. Introduction

Argument skills are fundamental skills for the 21st century citizens of democratic societies.

Argument skills are related with critical thinking (Kuhn, 2018; 2019), students' understanding of the nature and complexity of knowledge and students' conceptual understanding (Zohar & Nemet, 2002). Supporting students' argument skills, such as the ability to engage in argument from evidence and critique others' reasoning, is part of the standards that teachers are called upon to implement (NGSS Lead States, 2013; National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010). Although we have high expectations from teachers to promote students' argument skills, little is known of teachers' ability to fulfill this expectation. The present work examines teachers' ability to support the development of students' argument skills, focusing in particular on teachers' ability to evaluate arguments. Evaluating arguments is an essential skill for teachers in order to be able to offer constructive feedback to students which will enable them to improve their argument skills. Appropriate evaluation and feedback provide information on what high-quality performance looks like and in what direction students should work in order to improve their performance (Hattie & Timperley, 2007). In addition, to get a better understanding of how teachers evaluate arguments we have examined whether factors such as teachers' own skills of constructing arguments and their perceptions of their ability to effectively teach argument skills are related to their skills of evaluating arguments. Finally, given that the issue of domain specificity of argument skills is still an open one (Schauble, 2018), we have examined whether teachers' skills of evaluating and constructing arguments are domain specific.

1.1. Teachers' ability to support the development of their students' argument skills.

Given the prominent role of teachers in the educational system, several researchers examined the role of teachers in promoting students' argument skills. In particular, McNeil and Krajcik (2009) examined what kind of educational practices teachers use to support their students' ability to write scientific

justifications and explanations. The results, based on observations, showed that only one teacher – of the six who took part in the study – was able to support students in developing an understanding that evidence and explanations are the most important part of a scientific argument. Martin, Coffin and North (2007) examining 38 students of history teachers found that only 3.4% of their arguments contained evidence and counterarguments, suggesting that students need greater support for developing their argument skills. Similar findings have been reported by Muringani and Mavunga (2009), who carried out an investigation in order to examine whether teachers cultivate the argument skills of their students. The researchers examined 20 written essays, from tests in the courses of History and English. Results showed that English language learners constructed poor arguments because the teachers did not provide enough opportunities for them to exercise the skill properly, while in history the results were better because teachers in this area emphasized the development of argument skills more.

Overall, the studies examining teachers' ability to support the development of students' argument skills, using classroom observations, showed that not all teachers offer opportunities to their students and support in order to develop students' argument skills.

1.2. Teachers' perceived self-efficacy for teaching argument skills effectively.

Self-efficacy, defined by Bandura (1997) as “beliefs in one’s capabilities to organize and execute the courses of action required to produce given attainments” (p. 3), has been found to be related with individuals' behaviour, such as the persistence they show when dealing with obstacles, and their achievements (Bandura, 1997). Teachers' perceived self-efficacy, defined as teachers' belief in their personal efficacy to motivate and promote learning, appear to affect their teaching approach and the level of academic progress their students achieve (Holzberger, Philipp & Kunter, 2013). Reviews showed that teachers' self-efficacy is positively related with teachers' practices related to “classroom quality” (Zee & Kooman, 2016) and their students' achievement (Klassen & Tze, 2014). Teachers with high self-efficacy beliefs have been found to show effective classroom management (Woolfolk, Rosoff, & Hoy, 1990), to deploy more innovative teaching methods (Guskey, 1988), to encourage student autonomy (Woolfolk et al., 1990) and to be more positive in their rating of students' improvement regarding achievement and

effort (Miller, Ramirez, & Murdock, 2017). Studies on teachers' perceived self-efficacy for teaching argument skills are scant. Audeniz and Ozdilek (2015) found that teachers showed quite high self-efficacy regarding their ability to teach argumentation and their self-efficacy was further improved after engaging in an intervention. Yet, there is no evidence, as far as we know, examining the relation between teachers' self-efficacy on teaching argument skills and their ability to teach argument skills, including their ability to offer feedback and evaluate arguments.

1.3. The present study

The present study aims to form a better understanding of the support that teachers offer to their students for the development of their argument skills. We examine teachers' ability to construct arguments, their ability to evaluate arguments and their perceptions of their ability to teach effectively argumentation (Research question 1). We then examine if teachers' ability to construct argument and their perceptions of their ability to teach argumentation predict their ability to evaluate arguments (Research question 2). We have hypothesized that teachers would exhibit high self-efficacy regarding their ability to teach argument skills effectively and that their perceptions about their self-efficacy of teaching would be related to their skills in evaluating argument skills.

Finally, in the present work we address the issue of domain specificity of argument skills, examining whether teachers' ability of constructing and evaluating arguments, as well as their perceptions of their ability to teach argumentation, are domain specific (Research question 3). There is an open debate in the literature about the relative importance and role of domain-specific knowledge compared to the role of domain-general knowledge in students' and teachers' successful engagement in reasoning (Stevens, Wineburg, Herrenkohl, & Bell, 2005). To address this issue, we have included in our sample teachers from two different disciplines, social-science education teachers, who teach literature and history, and physical science education teachers, who teach science education. We have hypothesized that there will be differences among the two groups, of social-science and physical science education teachers, in how they use evidence in arguments in the science and history domains. In particular, we expect that teachers will

show better argument skills in their domain of expertise. To exclude the possibility that any possible differences between the two groups of teachers in a particular domain might be attributed to differences in their prior knowledge, we have also included a test of prior knowledge.

To examine teachers' ability to evaluate arguments and how this ability is related to teachers' own argument skills and their perceptions of their ability to teach argumentation skills effectively, we asked 104 secondary school teachers, social-science and physical science education teachers, to write an essay on a physical science topic and on a history topic in order to assess their argument skills in their domain of expertise and in another domain. Moreover, teachers were asked to rank 12 arguments as weak, moderate and good to assess their ability to evaluate arguments and to complete the Argumentation Self-Efficacy Scale (Audeniz & Ozdilek, 2015) to assess their perceptions of their ability to teach argumentation effectively.

2. Methods

2.1. Participants and design

Stratified sampling and then a simple random sampling were used in order to select a representative sample of 100 teachers out of the 302 secondary education teachers of a particular district in a South European country. We sent 100 questionnaires, 60 to social-science education teachers and 40 to physical science education teachers. 52 of the teachers who received the questionnaires completed them and returned them to the researcher. We also recruited a convenient sample of 52 secondary education teachers, from the same district, who attended a compulsory one-day professional development programme, organized annually by the Ministry of Education.

Overall, the participants of the study were 104 secondary education teachers, teaching 13- to 18-year-old students. Almost half of the participants, 53.8%, were social-science education teachers who teach literature or history ($M_{Age} = 35$, $SD = .92$; Females = 48), and the other half, 46.2%, were physical science education teachers who teach physical science ($M_{Age} = 37$, $SD = .75$; Females = 32).

2.2. Materials

2.2.1. Individual argument skills.

In order to evaluate the argument skills of teachers and especially their ability to use evidence in order to express their views on controversial topics, we asked the participants to read two texts – one a history issue and the other a socio-scientific issue – they were then asked to write a report on these two controversial issues.

The historical issue – adapted from the books written by Svolopoulos (2009) and Tsintilonis (2000) – was about a war from the recent history of participants' country. In particular, the text outlined a debate on whether the result of the war was the fault of their country's leader or not. The socio-scientific issue – adapted from the Allen and Freeman's (2002) article – was about the Kyoto Protocol. In particular, the text outlined the debate of whether or not the USA must sign the agreement.

After the teachers read the two texts, they were asked to write their opinion on the topic. In particular, they had to address the following tasks: "After reading the text, write your personal position about whether the outcome of the war was the fault of your country's leader or not" and, "After reading the text, write your personal position about whether the Kyoto Protocol should be signed by the USA or not."

2.2.2. Teachers' prior knowledge.

In order to investigate the prior knowledge of teachers, we asked the participants before completing the individual argument task, to answer two open-ended questions, concerning each of the issues that were used in the present study. The question for the historical issue was: "What do you know about the issue of accountability on the XXX war?"; and the question for the socio-scientific issue was: "What do you know about the issue of ratifying the Kyoto Protocol relating to the global planet?"

2.2.3. Teachers' ability to evaluate written arguments.

In order to investigate the ability of teachers to evaluate arguments we asked participants to rank 12 arguments as weak, moderate and strong. The first six arguments were about the history issue and the other six were about the socio-scientific issue. This tool format was based on other tools which were used in previous studies (Gwee & Damodaran, 2015). For each issue, we developed 2 weak arguments, 2 arguments of moderate quality and 2 strong arguments. Examples of each category of argument are

presented in Appendix A, Table A1. The results of our analyses, using Cronbach's alpha to estimate the internal consistency reliability, showed that the Scale's reliability was acceptable, for the history issue items ($\alpha = 0.80$) and for the socio-scientific issue items ($\alpha = 0.77$). Also, the correlation between the two sets of items was statistically positive ($r = 0.35$).

2.2.4. Teachers' perceptions about their ability to teach argument skills effectively.

In order to investigate teachers' perceptions about their ability to teach argument skills effectively as part of their educational practice, we used the Argumentation Self-efficacy Scale which was developed by Aydeniz and Ozdilek (2015). The Argumentation Self-efficacy scale consisted of three parts and included 30 items. Six items of the scale focused on argument development (planning) efficacy (e.g. I feel confident about my ability to design learning tasks to engage my students in small-group argumentation), 19 items were related to the lesson implementation (instruction) efficacy (e.g. I feel confident about my content knowledge to guide my students to effectively engage in argumentation) and five items were related to the assessment (evaluation) of students' learning during argumentation (e.g. I do not know how to use students' arguments as a measure of their learning). The rating range for all of the questions was from strongly disagree (score 1) to strongly agree (score 5) and was presented in a 1–5 Likert scale (1 = strongly disagree, 2 = disagree, 3 = undecided, 4 = agree, and 5 = strongly agree). To test the reliability of the scale, we calculated Cronbach's alpha. The results of our analyses showed that the Scale's reliability was good, for the first subscale ($\alpha = 0.76$), the second subscale ($\alpha = 0.80$), the third subscale ($\alpha = 0.77$) and the whole scale ($\alpha = 0.97$). The correlations between the subscales were statistically significant ($r = 0.49$ between the first and the second sub-scales; $r = 0.50$ between the first and the third sub-scales, and $r = 0.47$ between the second and the third sub-scales).

2.3. Procedure

Half of the participants completed a hard copy of the instruments in a room where the researcher was present, before the start of a professional development course. The other half of the participants completed the instrument online, using the Google Doc Forms tool. The tool controlled the time for completion of the instruments and the sequence of administration of the instruments. Teachers were asked

to first complete the prior-knowledge test, then the individual argument instrument and finally the test which assessed their perceptions about the importance of supporting the development of students' argument skills. The order of administration of the instruments on the two topics – history topic and socio-scientific topic – was counterbalanced across subjects. The instructions and the time that participants had available for completion of the instruments were identical between the participants who completed a hard copy of the instruments and the participants who completed the instruments electronically. No significant differences were observed between the participants who completed the instruments electronically and the participants who completed a hard copy of the instruments in their performance in all the variables that we examined in the present study. Hence, for simplicity, results for the two samples are combined.

2.4. Coding

2.4.1. Individual argument.

The length of the participants' responses in the individual arguments instrument ranged from 40 to 60 words in the history topic and from 40 to 50 in the socio-scientific topic for the social science education teachers and from 30 to 45 words in the history topic and from 35 to 50 in the socio-scientific topic, respectively, for the physical science education teachers. The social science education teachers' texts contained on average 3.13 idea units ($SD = 1.23$) in the history topic and 2.36 ($SD = 1.25$) in the socio-scientific topic and the physical science education teachers' text contained 2.71 units ($SD = .89$) in the history topic and 4.17 ($SD = 3.45$) in the socio-scientific topic. To examine the participants' argument skills, we employed three criteria. The first criterion (level of argument), examined whether participants focused only on their own position, if they criticized the opposing position, and whether they employed evidence. Responses to the initial questions in the individual argument skill instrument provided the basis for the coding of the participants' quality of argument. Firstly, the participants' responses were segmented into idea units. Each phrase which contained a new idea, position or statement was considered as a separate idea unit. Then each idea unit was ranked in one of five levels based on the coding schemes developed by Kuhn, Zillmer, Crowel and Zavala (2013) and Sadler (2006), based on Toulmin's (1958) work. (1. Claim without a justification, 2. Claim with justification, 3. Claim with justification and evidence, 4. Claim with

justification and evidence and a reference to the opposite position, 5. Claim with justification and evidence and a reference to the opposite position with evidence). A description of each level, along with examples are presented in Appendix A in Table A2. We focused particularly on Levels 4 and 5, which represent advanced argument skills.

The second criterion was the type of evidence employed, examining three types of evidence (a) evidence from the texts provided, (b) evidence from the teachers' own general knowledge, involving information that was not included in the text, and (c) evidence based on a combination of the two previous sources.

Finally, the third criterion was the amount of evidence that the participants employed in their arguments. We calculated the total number of pieces of evidence that participants employed in their report to address the question in the individual argument instrument, adding all the pieces of evidence that were included in all the idea units of participants' reports. The correlations between the criteria were statistically significant ($r = 0.35$ between the first and the second criteria; $r = 0.49$, between the first and the third criteria and $r = 0.52$, between the second and the third criteria).

Two coders – the first author and one other researcher – engaged in the coding of the participants' argument skills, blind to the identity and the specialization of the participants. Kappa statistic and the intraclass correlation co-efficient (ICC) were used to test interrater reliability. Reliability was calculated based on 25% of the data with Cohen's Kappa 0.84 for segmentation to idea units and 0.82 for judging the quality (level) of each idea unit. For the segmentation of the idea units the intraclass correlation co-efficient (ICC) was 0.942 and for coding the quality of idea units the ICC was 0.974 indicating good reliability. Any differences were resolved through discussion. After establishing inter-rater reliability, the remaining individual written texts were segmented and coded by one of the raters, the first author, again blind to the identity and the specialization of the participants.

2.4.2. Prior knowledge.

To code the participants' responses to the open-ended questions that were used to assess their prior knowledge on the history and socio-scientific topics, we developed a coding scheme based on existing

coding schemes for coding prior knowledge (Morrison & Luttenegger, 2015). In particular, the coding scheme involved four levels: The responses of the participants who claimed that they did not know anything about the issues and they could not provide any information or comment, were classified at the first level; the responses of the participants that claimed to have at least some relevant knowledge about the issues were classified at the second level; the responses of the participants who claimed that they were aware of the issues and were also able to make a simplified reference to it were classified at the third level; and the responses of the participants that claimed to have knowledge on the issue and made a detailed analysis of it were classified at the fourth level. The participants' responses to the prior knowledge task ranged from 60 words to 150 words in both studies. Cohen's kappa for assessing interrater reliability in coding the prior-knowledge test was 0.88.

2.4.3. Teachers' ability to evaluate written arguments.

Teachers' ability to evaluate individual written arguments was based on their judgments for ranking 12 arguments, 6 for each topic, as weak, of moderate quality, or as strong arguments. Each correct judgment received 1 point. Participants with 1 or 2 correct answers out of 6 were considered as not having a sufficiently developed ability to evaluate arguments. Participants with 3 or 4 correct answers out of 6 were considered as having developed a moderate capability for argument evaluation whereas participants with 5 or 6 correct answers out of 6 were considered as having a strong capability to evaluate arguments (Regala-Flores, 2007). For each participant, we calculated two scores, representing their capacity to evaluate arguments in the history topic and in the socio-scientific topic. Also, an overall score was calculated, by adding their scores on the two topics, representing their overall ability for argument evaluation. The results of our analyses showed that the Scale's reliability was acceptable, for the first history topic ($\alpha = 0.86$) and for the socio-scientific topic ($\alpha = 0.88$).

2.4.4. Teachers' perceptions about their ability to teach argument skills effectively.

Teachers' perceptions about their ability to teach argument skills effectively was judged based on their rating of the 30 questions of the Argumentation Self-efficacy Scale, which was developed by Aydeniz and Ozdilek (2015). The instrument included three aspects of teachers' perceptions about their effectiveness to

teach argumentation: (a) One aspect assessed the teachers' perceptions about their ability to plan their lesson, in order to develop students' argument skills (b) another aspect assessed the teachers' perceptions about their ability to implement a plan in the classroom and (c) the final aspect assessed the teachers' perceptions of their ability to assess the level of students' learning, concerning the construction of evidence-based arguments. Participants were asked to express the degree to which they agree with each statement presented in the instrument using a 1-5 Likert scale with rating range from strongly disagree (score 1) to strongly agree (score 5). A mean score was calculated for each of the three aspects of teachers' perceptions of their ability to teach argumentation that the instrument examined.

3. Results

3.1. Preliminary Analysis

Data were first screened for normality, to determine if they were normally distributed. All the variables were normally distributed and presented in the Table 1 below.

Insert Table 1

3.2.1. Teachers' skills for constructing arguments

Teachers' ability to construct arguments, to evaluate arguments and their perceptions of their ability to teach effectively argumentation were examined to address research question 1. To assess participants' skills for constructing arguments, we examined the participants' responses in the individual argument skill instrument. We employed three criteria to assess argument skills, the level of arguments, the type of evidence and the amount of evidence.

Level of argument. The results showed that in the history topic, 5.9% of the idea units produced by teachers – both social science and physical science education teachers – involved only claims, containing no data (Level 1), 23.1% included claims and justification (Level 2), 46% involved claims with justification and evidence (Level 3), 13.4% involved justification of the claim using evidence and a simple

reference to the opposing position (Level 4) and the 11.6% involved a justification of the claim by using evidence and a justification of the criticism of the opposing position by using evidence (Level 5). Our results showed that only a minority of teachers showed advanced argument skills, since only 13.4% of the arguments produced by teachers were of level 4 and even less, 11.6%, were of level 5. In the socio-scientific topic, 4.2% of the idea units involved only claims, containing no data (Level 1), 17.3% included claims and justification (Level 2), 60.2% involved claims with justification and evidence (Level 3), 11.6% involved justification of the claim using evidence and a simple reference to the opposing position (Level 4) and the 5.7% involved a justification of the claim by using evidence and a justification of the criticism of the opposing position by using evidence (Level 5). Our results showed that only a minority of teachers showed advanced argument skills, since only 11.6% of the arguments produced were of level 4 and even less, 5.7%, were of level 5.

Type of evidence used. We have also investigated the type of data that was used by the teachers both in the history and socio-scientific issues. The 64.3% of the participants used evidence from the texts provided by the researchers, 13.5% used evidence based on teachers' own general knowledge and 22.2% used evidence which was based on a combination of the two previous sources – general knowledge and the provided texts.

Amount of evidence used. Additionally, we investigated the amount of evidence used by teachers overall, in both topics. In the history topic, 26.8% of the teachers used one piece of evidence, 25% used two pieces of evidence, 9.6% used three pieces evidence, 2.9% used four pieces of evidence, and 1.9% used five pieces of evidence, 2.9% used eight pieces of evidence and 30.8% of teachers used no evidence. In the socio-scientific issue, 32.7% used one piece of evidence, 30.8% used three pieces of evidence, 3.8% used four pieces of evidence and 32.7% used no evidence.

3.2.2. Teachers' skill for constructing arguments in different domains

To examine whether the teachers' argument skills were specific to their domain of expertise (Research question 3), we examined social science and physical science education teachers' argument skills in the two topics.

We conducted an ANCOVA test with the quality of arguments, that is, the number of level 5 arguments produced, as the dependent variable, the specialization of teachers (social science and physical science education teachers) as the independent variable and prior-knowledge as the covariate. The results showed that there was no statistically significant difference between social science and physical science education teachers in the type of arguments they produced, neither in the history topic $F(1, 26) = 2.16, p = .13 - M = 1.04 (SD = 0.037)$ for the social science education teachers and $M = 1.1 (SD = 0.06)$ for the physical science education teachers – nor in the socio-scientific topic, $F(1, 26) = 2.14, p = .92 - M = 1.03 (SD = 0.17)$ for the social science education teachers and $M = 1.15 (SD = 0.17)$ for the physical science education teachers. Also, we conducted an ANCOVA test, including both level 4 and level 5 arguments as indicators for the quality of arguments for the dependent variable, the specialization of teachers (social science and physical science education teachers) as the independent variable and prior-knowledge as the covariate. The results showed that there was no statistically significant difference neither in the history topic $F(1, 25) = 1.24, p = .37$ nor in the socio scientific topic, $F(1, 23) = 1.27, p = .35$.

Type of evidence used. To examine further whether teachers' argument skills were specific to their domain of expertise we examined whether there were differences in the type of data that the teachers employed. We conducted an ANCOVA test with the evidence type as the dependent variable, the specialization of teachers (social science and physical science education teachers) as the independent variable and prior-knowledge as the covariate. The results for the history topic showed that social science education teachers used more evidence from a combination of general knowledge and the provided texts ($M = 1.6, SD = 0.07$) than physical science education teachers ($M = 1, SD = 0.04$), $F(1, 47) = 7.25, p = .01$. The results for the socio-scientific topic also showed a significant difference between the social science and physical science education teachers in the type of evidence they used, $F(1, 47) = 7.22, p = .01$. In particular, the physical science education teachers used more evidence from provided texts ($M = 2.8, SD = 0.06$) than the social science education teachers ($M = 2, SD = 0.04$).

Amount of evidence used. We conducted an ANCOVA test with the amount of evidence used by teachers as the dependent variable, the specialization of teachers (social science and physical science

education teachers) as the independent variable and prior-knowledge as the covariate, for the history and socio-scientific topics, to examine whether there were differences in the amount of the data used between the two groups of teachers. The results showed that there were no significant differences between the two groups of teachers on the total amount of evidence used by the participating teachers in neither the history topic, $F(1, 47) = 1.24, p = .28$ nor the socio-scientific topic, $F(1, 49) = 1.24, p = .24$.

3.3. Teachers' ability to evaluate written arguments in different domains

To assess teachers' ability to evaluate arguments, we examined their ability to evaluate arguments in (a) the history issue and (b) the social-scientific issue by asking them to rank 12 arguments, 6 for each topic, as weak, of moderate quality, or as strong arguments.

Social-science education teachers. On examining the social-science education teachers' argument evaluation ability in history, their domain of expertise, results showed that 12.6% of social science education teachers exhibited strong, 56.8% moderate and 30.6% weak argument evaluation skills. Regarding the social science education teachers' argument evaluation capacity in the socio-scientific domain, one that they were not experts in, results showed that 13.1% of social science education teachers showed good, 50.9% moderate and the 36% weak argument evaluation skills. Results showed that the majority of social science education teachers in both issues showed moderate evaluation capacity.

Physical science education teachers. Examining the physical science education teachers' argument evaluation skills in the history domain, one that they were not experts in, results showed that 18% of the physical science education teachers exhibited good, 46% moderate and 36% weak argument evaluation skills. With regards to the physical science education teachers' argument evaluation ability in their domain of expertise, on the socio-scientific issue, results showed that 14% of the physical science education teachers showed good, 36% moderate and 50% weak argument evaluation skills. Half of physical science education teachers showed moderate evaluation ability in the history issue and weak evaluation ability in the socio-scientific issue.

To examine if teachers' ability to evaluate arguments is domain specific (Research question 3), we compared teachers' ability to evaluate arguments in their domain of expertise and in another domain. A t-

test for dependent samples was conducted to compare the teachers' ability to evaluate the quality of arguments in their domain of expertise and in another domain. The analysis showed that there was no statistically significant difference in social science education teachers' ability to evaluate the quality of arguments in the history topic ($M = 3.5$, $SD = 1.4$) and in the socio-scientific topic ($M = 2.7$, $SD = 1.6$), $t(26) = 10$, $p = .108$. No significant differences were also observed in physical science education teachers' ability of evaluating the quality of arguments in the history ($M = 3.4$, $SD = 1.4$) and in the socio-scientific topic ($M = 2.8$, $SD = 1.4$), $t(26) = 1.07$, $p = .4$. The results showed that teachers' ability to evaluate arguments is not domain specific.

3.4. Teachers' perceptions about their ability to teach argumentation skills effectively

To examine teachers' perceptions about their ability to teach argumentation skills effectively during their educational practices, we examined the participants' responses in the Argumentation Self-Efficacy Scale. As seen in Table 2, the majority of social science and physical science teachers, 69%, agree that their lesson plan should include activities for the development of students' evidence-based argument skills and 69.2% of them believe that they have the proper skills to be able to evaluate their students' argument skills.

Insert Table 2

We further examined teachers' perceptions on the three sub-categories of the Argumentation Self-Efficacy Scale, that is the teachers' perceptions of their ability (a) to plan their lesson in order to develop students' argument skills (b) to implement a plan in the classroom and (c) to assess the level of students' learning, concerning the construction of evidence-based arguments. Table 3 shows the means of social science and physical science education teachers' perceptions of the plan learning activities, the implementation and the evaluation for the two topics.

To examine if there were domain differences in teachers' perceptions of their ability to teach argumentation (Research question 3) we compared teachers' perceptions in the two topics. T-tests showed that neither social science nor physical science education teachers' perceptions of their ability to teach

argumentation differ between the history and the socio-scientific topic ($t(26) = 3.32, p = .9$; $t(26) = 0.87, p = .6$, respectively). As seen in Table 3, teachers exhibited the same level of self-efficacy in teaching argumentation in both their domain of expertise and in another domain, outside of their expertise.

Table 3

3.5. Predicting teachers' ability to evaluate arguments from teachers' skill to construct argument and their perceptions of their ability to teach argumentation effectively

To examine our second research question, namely whether teachers' skills of constructing arguments and their perceptions about their ability to teach argument skills effectively, predicted teachers' ability to evaluate argument skills, we pursued a linear regression analysis, where the teachers' overall ability to evaluate arguments in both the history and the socio-scientific topic was the dependent variable and the teachers' overall argument skills in both the history and the socio-scientific issues, which were based on the quality of arguments – the number of level 5 arguments produced – and their overall perceptions about their ability to teach argumentation in both the history and the socio-scientific issue, were the predictor variables. Although the regression model was not statistically significant ($R^2 = .109, F(2, 24) = 2.59, p = .09$), teachers' skills to construct arguments was a statistically significant individual predictor ($B = 2.7, SE B = 1.19, b = .42, p = .03$), whereas teachers' perceptions of their ability to teach argumentation was not a statistically significant predictor ($B = -.02, SE B = 1.71, b = -.003, p = .98$). To examine further our findings, we run the model with only teachers' overall argument skills to construct arguments as the predictor and teachers' overall ability to evaluate arguments as the dependent variable. The regression model was statistically significant, $R^2 = .836, F(2, 163) = 4.36, p = .03$, and teachers' skills to construct arguments was a statistically significant individual predictor ($B = 2.9, SE B = 1.17, b = .42, p = .02$).

4. Discussion

The present study extends previous research by examining in-service teachers' ability to construct arguments, their ability to evaluate arguments and their perceptions about their ability to teach argument

skills effectively. It particularly examines whether teachers' ability to construct arguments and their perceptions of their ability to teach argument skills effectively predicts teachers' ability to evaluate arguments, which is fundamental for providing feedback and supporting students to sufficiently develop their argument skills. In addition, the present work examines if these skills – argument construction and evaluation – and perceptions of teaching argumentation, as well as the relation between them, are domain specific.

4.1. Teachers' ability to construct arguments, evaluate arguments and their perceptions of their ability to teach argumentation effectively

Our results show that teachers, both social science education teachers – who teach literature and history – and physical science education teachers – who teach science, did not have well-developed skills in neither the ability to construct arguments nor to evaluate arguments (Research question 1). In particular, teachers exhibited limitations in generating two-sided arguments and employing evidence systematically to justify their claims. Although teachers had university level education and they acquired **content** knowledge about their domain of expertise, they did not use evidence to justify their positions. The teachers appeared not to have sufficiently mastered the skill of constructing evidence-based arguments in order to make their arguments less vulnerable to a possible refutation. Similar conclusions were drawn by Cetin et al. (2014) and Simon et al. (2006), who found that pre-service and in-service teachers were unable to deal with controversial issues as they exhibited problems in producing counterarguments and rebuttal arguments. The fact that a considerable percentage of the teachers in the study expressed their opinions without justification, or they used claims in a simplistic way without using evidence, shows that they have not sufficiently developed the ability to construct evidence-based arguments and the understanding of the need to use evidence to back up a claim.

With regard to the teachers' ability to evaluate arguments, results show that the majority of social science and physical science education teachers had moderate to weak ability to evaluate arguments in both the history and the socio-scientific issue. This finding is consistent with Osborne et al. (2004), who also identified problems in the science educators' ability to evaluate students' arguments. The finding that

teachers exhibit difficulties in evaluating arguments is a matter of concern, given that appropriate evaluation and feedback is fundamental for supporting students to improve their performance (Hattie & Timperley, 2007). Given the central role of teachers in supporting students to develop their argument skills, special attention needs to be paid on how we can support teachers to develop their ability to evaluate arguments. Empirical evidence show that teachers benefited from training which aimed to develop their argument skills (Iordanou & Constantinou, 2014). The findings of the present study show that there is a need for programs tailored directly towards supporting all teachers, across all domains, to develop their argument skills and their understanding of the criteria of evaluating arguments. The content of those programs, whether for example they should focus on promoting epistemic norms, genre of exposition writing, or other futures in order to be effective in promoting teachers' argument skills, only further research can determine.

The finding of teachers' high perceived self-efficacy in teaching effectively argumentation in their classrooms, despite the fact that teachers' skills to construct and evaluate arguments were not sufficiently developed, is a matter of concern. Teachers reported that they could prepare and conduct lessons in such way as to support students' argument skills, particularly their ability to use evidence-based arguments. They also reported that they had the skills to assess effectively their students' arguments. Furthermore, teachers reported that they considered themselves capable to teach argumentation in the science domain to the same extent as they can teach argumentation in the history domain, despite their domain of expertise. This last finding may suggest that teachers' do not have a sufficiently developed awareness of the distinctive characteristics of the epistemology of each discipline, which is fundamental for developing learning goals and support the learning process of their students (see Goldman, Britt et al. 2016).

4.2. Domain-specificity of argument skills

Our results to the question of whether secondary school teachers' argument skills were specific to their domain of expertise (Research question 3), show that the teachers' argument skills were not specific to their domain of expertise. The teachers' prior knowledge did not seem to help in developing two-sided positions about an issue, as it was evident by the small percentage of teachers who produced

counterarguments even in their domain of expertise. These findings are in line with previous research (Iordanou & Constantinou, 2015), suggesting that prior knowledge is not a sufficient condition for skilled argumentation.

Some domain differences were observed regarding the type of evidence employed between the social science and physical science education teachers. Physical science education teachers used more evidence from the provided texts in the socio-scientific topic, whereas social science education teachers used a combination of evidence, from both their prior general knowledge and the provided texts, for the history topic. One possible explanation is that physical science education teachers have more objective views about knowledge and that is why they used only “objective” data whereas social science education teachers may believe that historical knowledge is more subjective, so that is why they also used their own opinions to support their views. This finding supports the view that the epistemic standards, that is the standards one employs for knowledge formation and evaluation, are domain specific (Duncan, Chinn, & Barzilai, 2018; Iordanou, 2016a; 2016b). The present data, though, does not enable us to draw conclusions regarding possible relations between teachers’ epistemic beliefs and their argument skills. Further research is required to examine this possibility. We cannot also rule out other possibilities which might have contributed to the results, such as emotions and identity (Darner, 2019; Nyhan & Reifler, 2010; Rosenau, 2012), given that the history topic was relevant to the participants’ country whereas the socio-scientific topic was about someone else’s country, which constitutes one of the limitations of the present study. Future research is indicated to examine further the issue of domain-specificity of teachers’ argument skills, using different topics.

Teachers’ ability to evaluate arguments and their perceptions of their ability to teach argumentation appeared also not to be domain specific. Although one would expect teachers to exhibit a better ability to evaluate arguments and a higher self-efficacy in teaching argumentation in their domain of expertise than in another domain, this was not the case. One possible explanation of the finding that no significant differences were observed in teachers’ ability of evaluating the quality of arguments between the history and the socio-scientific topic, might be teachers’ argument skills, which have been found in the present

study not to be sufficiently developed in neither their domain of expertise nor in another domain outside of their expertise. This possibility is further examined in the section 4.3. below.

4.3. Predicting teachers' ability to evaluate arguments from their ability to construct arguments and their perceptions of their ability to teach argument skills

On examining if the teachers' ability to evaluate arguments and their perceptions of their ability to teach argumentation predicted teachers' ability to evaluate arguments (Research question 2), our findings showed that the teachers' ability to construct arguments predicted their ability to evaluate arguments. This finding is in line with our hypothesis, showing that if teachers do not have strong argument skills how will they be able to evaluate their students' arguments and support them to develop this ability. This finding has important educational implications, suggesting that investing in promoting teachers' own argument skills might be a promising endeavor towards equipping teachers with the skills needed in order to support their students in developing argument skills.

The teachers' perceptions of their ability to teach argumentation effectively – perceived self-efficacy – did not predict their ability to evaluate arguments. In general, teachers appeared to be very confident regarding their ability to teach argumentation. In other words, the teachers, regardless of the level of their argument skill, considered themselves capable of supporting their students to develop strong argument skills. This finding shows that there is a gap between teachers' perceptions of their ability to teach argument skills effectively and their actual skills in constructing and evaluating arguments. Teachers seem to over-estimate their ability to teach argumentation or do not have self-awareness of their argument skills and their ability to evaluate arguments. Future research needs to examine how teachers build their teaching self-efficacy (Morris, Usher & Chen, 2017), particularly their self-efficacy for teaching argumentation, and how to support teachers to develop better self-awareness of their argument skills.

4.4. Educational Implications

The findings of the present study, showing that teachers do not have adequately developed argument skills and that they exhibit difficulties in assessing arguments, are of particular concern. Since the

development of students' argument skills constitutes one of the main goals of contemporary curriculum (NGSS Lead States, 2013; National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010), it is important to equip teachers with the necessary tools that will enable them to support their students in achieving this goal. Yet, the finding of the present study, of a connection between teachers' ability to evaluate arguments and their own skills of constructing arguments, has important educational implications, suggesting that teachers' own argument skills play a key role in their ability to evaluate arguments which is closely connected with teachers' ability to support the development of students' argument skills. Teachers need to develop their own argument skills in order to be able to help their students develop theirs, by engaging in particular programs which aim to develop argument skills (Iordanou & Kuhn, 2019; Iordanou, Kuhn, Matos, Shi & Hemberger, 2019).

Education of the 21st century should aim to support the development of thoughtful, active and emancipated students, with strong critical thinking skills, who will be able to respond effectively to the various challenges of modern times. Therefore, reconsidering the educational practices that are currently employed is necessary and crucial in order to be confident that we provide students with all the necessary knowledge and skills, with which they could strengthen the democratic nature of human society. Now that the phenomenon of globalization is expanding, education must support the development of students' argument skills, so that they will be able, as responsible citizens, to take positions on critical, local and global issues and to create a fairer society.

References

- Allen, B. C., & Freeman, H. C. (2002). The petition: A global warming case study. In *National Center for Case-study Teaching in Science*. Retrieved November 20 2017, from http://www.sciencecases.org/petition/petition_notes.asp
- Aydeniz, M., & Ozdilek, Z. (2015). Assessing and enhancing pre-service science teachers' self-efficacy to teach science through argumentation: challenges and possible solutions. *International Journal of Science and Mathematics Education, 13*(3), 1-19.
- Avraamidou, L., & Zembal-Saul, C. (2005). Giving priority to evidence in science teaching: a first-year elementary teacher's specialized practices and knowledge. *Journal of Research in Science Teaching, 42*(9), 965-986.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.
- Billig, M. (1989). *Arguing and thinking: A rhetorical approach to social psychology*. Cambridge: Cambridge University Press.
- Darner, R. (2019). How can educators confront science denial?. *Educational Researcher, 48*(4), 229-238.
- Darling-Hammond L. (2000). Teacher quality and student achievement: A review of state policy evidence. *Education Policy Analysis Archives, 8*(1).
- Duncan, N. (2007). Feed-forward: improving students' use of tutor comments. *Assessment & Evaluation in Higher Education, 32*(3), 271 -283.
- Duncan, R. G., Chinn, C. A., & Barzilai, S. (2018). Grasp of evidence: Problematizing and expanding the next generation science standards' conceptualization of evidence. *Journal of Research in Science Teaching, 55*(7), 907–937.
- Erduran, S., Simon, S., & Osborne, J. (2004). TAPing in argumentation: Developments in the application of Toulmin's argument pattern for studying science discourse. *Science Education, 88*(6), 915-933.
- Goldman, S. R., Snow, C., & Vaughn, S. (2016). Common themes in teaching reading for understanding: Lessons from three projects. *Journal of Adolescent & Adult Literacy, 60*(3), 255-264.

- Guskey, T., R. (1988). Teacher efficacy, self-concept, and attitudes toward the implementation of instructional innovation. *Teaching and Teacher Education*, 4(1), 63-69.
- Gwee, S., Damodaran, S. (2015, October). *Use of Web 2.0 and mobile technologies for developing argumentative skills*. Paper presented at the 14th World Conference on mobile and contextual learning, Venice, Italy.
- Hattie, J., & Timperley, H. (2007). The Power of feedback. *Review of Educational Research*, 77, 81-112.
- Holzberger, D., Philipp, A., & Kunter, M. (2013). How teachers' self-efficacy is related to instructional quality: A longitudinal analysis. *Journal of Educational Psychology*, 105(3), 774-786.
- Iordanou, K. (2016a). From theory of mind to epistemic cognition. A Lifespan perspective. *Frontline Learning Research*, 4(5), 106 – 119.
- Iordanou, K. (2016b). Developing epistemological understanding through argumentation in scientific and social domains. *Zeitschrift für Pädagogische Psychologie*. 30(2-3), 109-119.
- Iordanou, K. & Constantinou. C. P. (2014). Developing pre-service teachers' evidence-based argumentation skills on socio-scientific issues. *Learning & Instruction*. 34, 42-57.
- Iordanou, K. & Constantinou. C. P. (2015). Supporting use of evidence in argumentation through practice in argumentation and reflection in the context of SOCRATES learning environment. *Science Education*, 99, 282–311.
- Iordanou, K., & Kuhn, D. (2019). Contemplating the Opposition: Does a Personal Touch Matter? *Discourse Processes*.
- Iordanou, K., Kuhn, D., Matos, F., Shi, Y., & Hemberger, L. (2019). Learning by arguing. *Learning and Instruction*, 63, 101-207.
- Jimenez-Aleixandre, M. P., & Erduran, S. (2007). Argumentation in science education: an overview. In S. Erduran & M. P. Jimenez-Aleixandre (Eds.), *Argumentation in Science Education* (pp. 3–27). The Netherlands: Springer.
- Klassen, R. M., & Tze, V. M. (2014). Teachers' self-efficacy, personality, and teaching effectiveness: a metaanalysis. *Educational Research Review*, 12, 59–76.

- Kuhn, D. (2018). *Building our best future: Thinking critically about ourselves and our world*. New York: Wessex Press.
- Kuhn, D. (2019). Critical Thinking as Discourse. *Human Development*, 62(3), 146-164.
- Kuhn, D., & Moore, W. (2015). Argumentation as core curriculum. *Learning: Research and Practice*, 1(1), 66-78.
- Leder, G., C., Pehkone, F. & Töner, G. (2002). *Beliefs: A hidden variable in Mathematics Education*. Klumer Academic Publishers, Netherlands.
- Martin, D., Coffin, C., & North, S., (2007). What's your claim? Developing pupils' historical argument skills using asynchronous text based computer conferencing. *Teaching History*, 126, 32-37.
- McNeill, K., & Krajcik, J. (2009). Synergy between teacher practices and curricular scaffolds to support students in using domain-specific and domain-general knowledge in writing arguments to explain phenomena. *The Journal of the Learning Sciences*, 18(3), 416-460.
- Mercer, N., Dawes, L., Wegerif, R., & Sams, C. (2004). Reasoning as a scientist: ways of helping children to use language to learn science. *British Educational Research Journal*, 30(3), 359-377.
- Miles, MB., & Huberman, A.M. (1994). *Qualitative Data Analysis* (2nd edition). Thousand Oaks, CA: Sage Publications.
- Morris, D. B., Usher, E. L., & Chen, J. A. (2017). Reconceptualizing the sources of teaching self-efficacy: A critical review of emerging literature. *Educational Psychology Review*, 29(4), 795-833.
- Morrison, A.D., & Lutteneger, K. C. (2015). Measuring Pedagogical Content Knowledge Using Multiple Points of Data. *The Qualitive Report*, 20(6), 804-816.
- Miller, A. D., Ramirez, E. M., & Murdock, T. B. (2017). The influence of teachers' self-efficacy on perceptions: Perceived teacher competence and respect and student effort and achievement. *Teaching and Teacher Education*, 64, 260-269.
- National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common Core State Standards*. Washington, DC: National Governors Association Center for Best Practices & Council of Chief State School Officers.

NGSS Lead States. (2013). *Next Generation Science Standards: For States, By States*. Washington, DC:

The National Academies Press.

Nyhan, B., & Reifler, J. (2010). When corrections fail: The persistence of political misperceptions.

Political Behavior, 32(2), 303-330.

Osborne, J. (2010). Arguing to Learn in Science: the role of collaborative, critical discourse. *Science*, 328

(5977), 463-466.

Osborne, J., Erduran, S., & Simon, S. (2004). Enhancing the quality of argumentation in school science.

Journal of Research in Science Teaching, 41(10), 994 – 1020.

Osborne, J. F., Simon, S., & Collins, S. (2003). Attitudes towards Science: A Review of the Literature and

its Implications. *International Journal of Science Education*, 25(9), 1049–1079.

Pajares, M. F. (1992). Teachers' beliefs and educational research: Cleaning up a messy construct. *Review*

of Educational research, 62(3), 307-332.

Plantin, C. (1996). *L'argumentation [Argumentation]*. Paris, France: Seuil.

Regala-Flores, E. (2007). Thinking skills reflected in the argumentative essays of freshman college

students: A descriptive analysis. *The Asia Pacific-Education Researcher*, 16(1), 33-44.

Roberts, R., & Gott, R. (2010). Questioning the evidence for a claim in a socio-scientific issue: an aspect of

scientific literacy. *Research in Science & Technological Education*, 28(3), 203-206.

Rosenau, J. (2012). Science denial: A guide for scientists. *Trends in microbiology*, 20(12), 567-569.

Ross, J. A., & Bruce, C. (2007). Self-assessment and professional growth: The case of a grade 8

mathematics teacher. *Teaching and Teacher Education*, 23, 146–159.

Sadler, T. D. (2006). Promoting Discourse and Argumentation in Science Teacher Education. *Journal of*

Science Teacher Education, 17(4), 323-346.

Schauble, L. (2018). In the eye of the beholder: Domain-General and domain-specific reasoning in science.

In *Scientific Reasoning and Argumentation* (pp. 21-43). Routledge.

Stevens, R., Wineburg, S., Herrenkohl, L. R., & Bell, P. (2005). The comparative understanding of school

subjects: Past, present, and future. *Review of Educational Research*, 75(2), 125-157.

- Svolopoulos, K. D. (2009). *The decision on the extension of Greek domination in Asia Minor*. National Research Foundation Eleftherios Venizelos, Icarus.
- Toulmin, S. E. (1958). *The uses of argument*. Cambridge: University Press.
- Tsintilonis, C. (2000). *Asia Minor Destruction-80 years of the Treaty of Sevres*. Communist Review.
- Walker, C. M., Wartenberg, T., & Winner, E. (2012). Examining the effects of philosophy classes on the early development of argumentation skills. In S. Goering, N. Shudak, & T. Wartenberg (Eds.), *Philosophy in schools: An introductory handbook for philosophers and teachers*, (pp. 277-287). Routledge Press: New York.
- Wecker, C., Hetmanek, A., & Fischer, F. (2014). The interplay of domain-specific and domain-general factors in scientific reasoning and argumentation. In J. L. Polman, E. A. Kyza, D. K. O'Neill, I. Tabak, W. R. Penuel, A. S. Jurow, K. O'Connor, T. Lee & L. D'Amico (Eds.), *International Conference of the Learning Sciences: Learning and Becoming in Practice* (Vol. III, pp. 1192–1194). Boulder, CO: International Society of the Learning Sciences.
- Woolfolk, A. E., Rosoff, B., & Hoy, W., K. (1990). Teachers' sense of efficacy and their beliefs about managing students. *Teaching and Teacher Education*, 6(2), 137-148.
- Zee, M., & Koomen, H. M. (2016). Teacher self-efficacy and its effects on classroom processes, student academic adjustment, and teacher well-being: A synthesis of 40 years of research. *Review of Educational research*, 86(4), 981-1015.
- Zohar, A., & Nemet, F. (2002). Fostering students' knowledge and argumentation skills through dilemmas in human genetics. *Journal of Research in Science Teaching*. 39(1), 35 – 62.

Table 1

Descriptive Statistics for Participants' Argument Skills, Ability to Evaluate Arguments, Perceptions of their Ability to Teach Argumentation and Prior Knowledge

Variables	Mean	SD	Skewness	Kurtosis
Teachers' skill for constructing arguments				
Mean number of arguments of level 4 and 5 constructed	2.6	.71	-.451	-.79
Mean of number of evidence from the text or from a combination of general knowledge and text used (Type of evidence)	3.1	1.41	-.073	-.733
Total amount of evidence in history issue	2.55	1.37	3.15	-1.59
Total amount of evidence in socio-scientific issue	4.93	3.67	.877	-1.74
Total number of evidence in history and Socio-scientific issue	2.76	.92	.004	-.899
Teachers' skill for evaluating arguments				
Social science education teachers' argument evaluation capability in history, based on the number of correct evaluation of arguments	3.56	1.33	.512	-.312
Physical science education teachers' argument evaluation capability in socio-scientific issue based on the number of correct evaluation of arguments	2.79	.95	-.362	-.486

TEACHERS ABILITY TO CONSTRUCT AND EVALUATE ARGUMENTS

Teachers' overall capability to evaluate arguments in history and socio-scientific issue based on the number of correct evaluation of arguments	6.34	2.98	.322	-.433
Score of teachers' perceptions of their ability to teach argumentation	3.82	.54	-.162	-.344
Mean of Prior-Knowledge in the history topic	2.87	.78	-.99	-.22
Mean of Prior-Knowledge in socio-scientific topic	2.47	.89	-.61	-.39

Table 2

Teachers' Perceptions about their Ability to Teach Argumentation Skills Effectively

Teachers' Perceptions	Disagree Completely	Disagree	Undecided	Agree	Agree Completely
Plan learning activities to engage students in evidence-based arguments	1.5%	6.9%	23%	33.7%	35.3%
Implement a plan in the classroom	3.3%	11.8%	22.3%	40.7%	19.95%
Evaluate students' evidence-based arguments	1.3%	11.2%	18.3%	44.9%	24.3%

Table 3

The Means of the Planned Learning Activities, the Implementation and the Evaluation for the two Groups for the Two Topics

Variables	Social-science education teachers in history	Social-science education teachers in socio-scientific issue	Physical science education teachers in history	Physical science education teachers in socio-scientific issue
Teachers' perceptions about their ability to:				
Plan learning	3.97 (<i>SD</i> = 0.58)	3.6 (<i>SD</i> = 0.46)	3.84 (<i>SD</i> = 0.63)	3.65 (<i>SD</i> = 0.48)
Implement the Plan	3.52 (<i>SD</i> = 0.6)	3.73 (<i>SD</i> = 0.46)	3.63 (<i>SD</i> = 0.46)	3.83 (<i>SD</i> = 0.76)
Assess students' learning	3.65 (<i>SD</i> = 0.69)	3.51 (<i>SD</i> = 0.35)	3.75 (<i>SD</i> = 0.56)	3.8 (<i>SD</i> = 0.65)

APPENDIX A

Table A1: Categories of Arguments

Categories of Arguments	Examples
1. Claim without evidence	<p>A1. It is certain that the outcome of the war was the country's leader fault.</p> <p>A2. USA should sign the Kyoto Protocol.</p>
2. Claim with evidence	<p>B1. The the outcome of the war was the country's leader fault because he was responsible for the deficiencies in the strategic planning of the campaign.</p> <p>B2. The USA should not sign the Kyoto Protocol as it would not help the greenhouse effect because the cycles of cooling and heating are alternated during the history of earth.</p>
3. Claim with a reference to the opposing position and employment of evidence	<p>C1. The political opponents of country's leader are responsible for the outcome of the war, because they marched recklessly into the area. But maybe it is also The country's leader fault, because he did not take into account the geographic characteristics of the area.</p>

Table A2: Levels of participants' arguments

Level	Example
1.Claim without a justification	The country's leader responsibilities were enormous for many reasons.
2.Claim with justification	USA should sign the Kyoto Treaty because it is a factor that could help to deal with the greenhouse effect.
3.Claim with justification and evidence	In my opinion, the outcome of the war was the fault of the leaders' opponents, because although they won over XXX in the elections, they continued the campaign into the country without studying a withdrawal plan and without assessing the data and only making decisions with political considerations.
4.Claim with justification and evidence and a reference to the opposite position	The Kyoto Protocol must be signed because the USA is one of the most polluting countries of the planet. The USA must help to tackle the greenhouse effect because the warming temperatures show dramatic changes while the ice melts. But perhaps the signing of the treaty is unnecessary because there are indirect ways that

- could make the treaty invalid.
5. Claim with justification and evidence and a reference to the opposite position with evidence
- I think that it was the country's leader fault because he made some diplomatic mistakes. He had not secured assurances on the allies' help, while he had to realize that the dream of the fulfillment of the Great Idea was a utopia. But also his political opponents had not handled correctly the political balances and as a result they lost the confidence of the Powers (powerful countries) who they always aimed in their own interests.
-