

Revealing argumentation skills through thought experiments in mechanics

Özge Tağluktimur¹, Gül Ünal Çoban² 

¹ Dusunur Private Schools, İzmir, Türkiye; ² Dokuz Eylül University, Buca Faculty of Education, Department of Science Education, İzmir, Türkiye

ABSTRACT

Thought experiments (TEs) have long been used in physics research and education, yet their role in science teacher education—particularly in enhancing prospective science teachers' (PSSTs) argumentation skills—remains underexplored. This study aimed to examine the effects of worksheets enriched with TEs on PSSTs' conceptual understanding of mechanics and their ability to construct scientific arguments. A pretest-posttest control group design was implemented with 35 first-year PSSTs enrolled in a Physics I Laboratory course. Participants were randomly assigned to either an experimental group (n=17) or a control group (n=18), and engaged in laboratory activities over four weeks, totaling 10 hours. Both groups followed similar worksheet-based instruction, with the key difference being the inclusion of TEs in the experimental group's materials. PSSTs in both groups conducted the same experiments in argumentation skills, suggesting that TEs can serve as effective tools in fostering scientific reasoning. These findings underscore the potential of integrating TEs into physics laboratory instruction as a means to enhance argumentation competence in science teacher education.

ARTICLE HISTORY

Received 18 June 2025

Accepted 15 December 2025

KEYWORDS

Argumentation skills
Thought experiments
Prospective science teachers
Physics teaching

CORRESPONDENCE

Gül Ünal Çoban



Introduction

The education of preservice science teachers (PSSTs) places significant emphasis on laboratories apart from textbook or lecture style instruction. The educational effectiveness of laboratories is an undeniable fact providing support for higher order learning skills that include observing, planning an experiment, asking relevant questions, hypothesizing, and analyzing experimental results (Bybee 2000; Hofstein et al. 2004). On the other hand, while many conceptual topics can be directly tested under laboratory conditions, certain subjects, such as special and general relativity, quantum mechanics, and scientific revolutions, pose challenges for hands-on experimentation. Additionally, topics like force and simple machines (e.g., inclined planes, pulleys, levers, gears) are easily observable and relatable, while idealized scenarios, such as frictionless environments, are harder to experience directly. To help students grasp these concepts, educators can utilize TEs in educational settings. TEs not only encourage mental visualization of the related concepts but also represent a rich laboratory for cognitive learning, incorporating intense conceptual and methodological processes without requiring practical application.

Learning in the laboratory also provide the PSSTs learn how to collect, use, and interpret data; make claims using evidence, and discuss science as debate to support claims with evidence from data (National Research Council, 2012). In other words, laboratory activities help individuals construct arguments and epistemic values of their knowledge claims. In terms of cognition, developing an argument is a conceptual process that helps in grasping aforementioned concepts. Additionally, reasoning skills, which entail linking claims to evidence, are strengthened through this process (Osborne, 2010).

While TEs have been widely recognized as valuable tools for fostering critical thinking and conceptual understanding in science education (Najami et al., 2020; Yalçın, 2018; Aktaş & Doğan, 2018; Altınok, 2017,

McDonald, 2014; Ozdem et al., 2013), there is a notable lack of research exploring their specific impact on argumentation skills. Existing studies have primarily focused on the role of TEs experiments in enhancing conceptual knowledge or problem-solving abilities, leaving their potential to develop reasoning and argumentation skills underexplored. Furthermore, little attention has been given to how TEs might influence the argumentation competencies of prospective science teachers, who play a critical role in modeling these skills for future generations. Addressing this gap is essential to better understand the pedagogical potential of TEs in teacher education and their broader implications for science learning. The objective of this study is to investigate how worksheets enriched with TEs impact PSTs' conceptual understanding in mechanics, and argumentative skills. Additionally, the study seeks to gather candidates' reflections on the activities involved in the process.

TEs

TEs are defined by Reiner and Gilbert (2000a) as "designs created to test the validity of a claim or to persuade others of that validity." Similarly, Sorensen (1992a) describes them as experiments emphasizing conceptual design rather than practical application. These experiments leverage pre-existing knowledge to construct hypothetical scenarios that either support or challenge a claim (Brendel, 2004). In physics education, TEs are particularly valuable as they allow students to engage with idealized conditions and mental simulations, fostering an understanding of natural laws without requiring physical experimentation.

Four key characteristics distinguish TEs:

1. **Cognitive Nature:** TEs occur in the "laboratory of the mind" (Brown, 1991) and function as cognitive exercises rather than physical processes.
2. **Narrative Structure:** They follow a structured narrative detailing the design and reasoning process (Nersessian, 1992a; Gooding, 1993, as cited in Ateş, 2015).
3. **Elimination of Technical Constraints:** They unfold in ideal conditions based purely on reasoning and existing knowledge, free from material limitations (Irvine, 1991).
4. **Scientific and Philosophical Inquiry:** Unlike literary storytelling, they are employed to explore scientific or philosophical questions, generating mental observations (Brown, 2001; Ateş, 2015).

As the above characteristics point out, TEs serve various educational purposes, such as supporting claims, generating evidence, testing theoretical consistency, and broadening conceptual perspectives (Brown, 2006). Kuhn (1977) also categorizes TEs in physics as critical, explanatory, or supportive. For instance, Einstein's "following a beam of light" experiment exemplifies a critical TE, while Newton's "bucket experiment" serves a constructive function, and Stevin's "chain experiment" aids in idealization (Ateş, 2015). Nersessian (1991) further highlights that constructing and executing TEs enables individuals to identify conceptual limitations in existing theories, reinforcing their role in conceptual change. Thus, TEs are essential in physics, facilitating the evolution of explanatory theories and the refinement of scientific concepts.

While TEs have been widely employed in physics research and education, their potential in science teacher education remains underexplored. PSTs must develop strong conceptual understanding and scientific reasoning skills, and incorporating TEs into their training can significantly enhance these competencies (Matthews, 2014). Research suggests that TEs can bridge the gap between theoretical knowledge and practical application, equipping PSTs with effective strategies for classroom instruction (Gilbert & Reiner, 2000).

TEs' contribution to teacher education can be grouped under four main areas. First, engaging with TEs supports the development of argumentation and critical thinking skills by encouraging pre-service teachers (PSTs) to construct logical arguments, evaluate assumptions, and refine their reasoning (Erduran & Jimenez-Aleixandre, 2008). In addition, TEs can enhance conceptual understanding by providing mental models that help PSTs and their future students clarify and internalize complex scientific ideas (Reiner and et al., 2000). When integrated into pedagogical training, TEs also support inquiry-based learning by enabling PSTs to facilitate discussions where students are encouraged to actively explore and question scientific concepts (Osborne and et al., 2004). Finally, TEs contribute to addressing misconceptions by offering a structured way to test claims and assumptions, which helps PSTs identify and correct students' misunderstandings more effectively (Matthews, 2014). These contributions highlight the multifaceted value of TEs in equipping future science teachers with the cognitive and pedagogical tools necessary for effective science instruction.

Despite these benefits, empirical research on the direct impact of TEs on PPSTs' pedagogical skills and argumentative reasoning is still limited (Gilbert & Reiner, 2000). Existing studies primarily focus on historical or theoretical aspects, rather than assessing their effectiveness in laboratory or classroom settings (Reiner et al., 2000). To address this gap, teacher education programs should incorporate explicit instruction on TEs as a pedagogical tool. Training PPSTs to reconstruct TEs as structured arguments (Norton, 1996) can enhance their ability to identify biases, clarify assumptions, and apply logical reasoning in teaching science. Reiner & Burko (2003) warn that TEs, like laboratory experiments, are susceptible to misinterpretation due to cognitive biases. However, by integrating Norton's argument-based approach, PPSTs can strengthen the epistemological foundations of TEs, making them more reliable and valid for classroom use.

While TEs are well-established tools in physics, their role in science teacher education remains underutilized. Research suggests that integrating TEs into teacher training can enhance conceptual learning, argumentation skills, and inquiry-based teaching. However, further empirical studies are needed to validate their effectiveness in improving pre-service teachers' pedagogical skills and students' scientific reasoning abilities. Future teacher training programs should emphasize structured TE-based instruction, equipping prospective teachers with powerful tools for fostering scientific literacy and conceptual understanding in the classroom.

Argumentation

Argumentation is a recommended framework in science education as it aids in students' development of scientific epistemology, conceptual understanding, scientific literacy (reading and writing in the language of science), and scientific reasoning skills (Erduran and et al., 2004; Jimenez-Aleixandre & Erduran, 2007). In research, various definitions of argumentation emerge; Kuhn (1991) describes it as a claim accompanied by reasoning, while Kaya & Kılıç (2008) define it as rational and reasonable discussions aimed at explaining the uncertainty between different situations.

One widely used framework for argumentation is the Toulmin argument model, which includes six subcomponents: claim, data, warrant, backing, rebuttal, and qualifier. Walton & Sartor (2013) describe the claim component as a statement reflecting an individual's thoughts or a viewpoint being proven. The data subcomponent consists of facts, evidence, and examples supporting the claim. The warrant explains the connection between data and claim through supporting statements. Backing consists of theoretical assumptions or justifications that validate the warrant, whereas the rebuttal refers to conditions or counterarguments that challenge the claim. These components help students develop reasoning skills by analyzing and organizing information to construct meaningful arguments. When teaching mechanical topics, it is crucial to determine the conditions under which a claim is valid and support it with justifiable data.

In the context of teacher education, argumentation plays a critical role in shaping PSSTs' ability to facilitate student-centered learning, encourage critical thinking, and promote scientific discourse in the classroom (Driver and et al., 2000). Studies indicate that PSSTs often struggle with integrating argumentation into their teaching due to a lack of experience and instructional strategies (McNeill & Pimentel, 2010). Engaging prospective teachers in argumentation-based instruction enhances their pedagogical content knowledge and ability to scaffold student reasoning (Osborne and et al., 2004). For instance, Sadler & Zeidler (2005) found that argumentation enables PSSTs to address socio-scientific issues more effectively, helping students navigate complex, real-world scientific problems. Additionally, Erduran & Jimenez-Aleixandre (2008) emphasize that structured argumentation activities improve teachers' understanding of the nature of science, fostering a deeper appreciation for evidence-based reasoning in scientific inquiry. Moreover, research highlights that teacher training programs should incorporate explicit instruction on argumentation to help PSSTs develop confidence in facilitating classroom discussions, assessing students' arguments, and addressing misconceptions effectively (Simon and et al., 2006). Studies suggest that when teachers receive argumentation-focused training, they are more likely to create inquiry-based learning environments and encourage students to construct and evaluate scientific explanations critically (Berland & McNeill, 2010).

In conclusion, integrating argumentation in science teacher education not only enhances prospective teachers' scientific reasoning and pedagogical skills but also prepares them to cultivate a classroom culture where students engage in evidence-based discussions, think critically, and develop deeper conceptual understandings.

Research goal and question

Incorporating TEs into classroom can significantly enhance students' understanding of complex concepts by allowing them to explore hypothetical scenarios and their implications without the need for physical

experiments as well as PSSTs. TEs encourage students to engage in abstract reasoning and critical thinking, which are crucial skills in the learning process. For instance, by imagining scenarios where physical constraints are removed or altered, students can better grasp theoretical principles and develop a deeper comprehension of the subject matter (Larkin, 2017). The use of TEs in the classroom also promotes active learning, as students actively engage with the material by constructing and analyzing mental models (Gilbert & Boulter, 2000). This approach not only aids in understanding but also fosters creativity and problem-solving skills, making it a valuable educational tool in enhancing student learning experiences. Students begin by analyzing the problem addressed by TE and then align the actual scenario with the theoretical model proposed. This alignment helps them apply the model effectively to solve the problem. During the TE phase, students are expected to reason by drawing analogies and using the dynamic, interactive nature of mental models to visualize and understand the relationships between various components of the model (Unal, 2009). Moreover, TEs in educational materials, such as worksheets, is expected to positively impact science teacher candidates' attitudes towards physics and laboratory courses, as well as improve their performance in mechanics. This integration can lead to a more profound understanding of physical concepts and enhance overall educational outcomes.

Mechanics, the oldest field in physics, studies the relationships between force, matter, and motion. TEs are often employed in physics to examine problems difficult to address through physical trials, making them invaluable in teaching mechanics (Gilbert & Boulter, 2000). Since abstract concepts and laws in mechanics may not be directly observable, TEs enable students to visualize these principles, bridging the gap between theory and mental experimentation (Gendler, 2007). Moreover, as mentioned previously, explicitly incorporating argumentation into educational settings where TEs are conducted can help reduce reliance on imagination and biases, guiding reasoning through structured argumentation. Especially in teaching mechanics, it is essential to examine the validity of claims about concepts, supporting these claims with evidence and reasoning.

Studies investigating argument skills have examined how socioscientific topics, argumentation-based teaching, and argumentation-focused lab activities influence students' argumentative skills (Najami et al., 2020; Yalçın, 2018; Aktaş & Doğan, 2018; Altınok, 2017; McDonald, 2014; Ozdem et al., 2013).

In the same manner, studies on implementation of TEs in educational settings reveal, that the researchers have examined students' cognitive processes (Acar, 2013) and explore the effects of TE-based arguments on critical thinking (Tüzün & Köseoğlu, 2018). However, there is limited research specifically on how TEs affect conceptual learning or argumentative skills of PSSTs (Cetinkaya, 2019; Cetin, 2014, Aydeniz & Dogan, 2016; Reiner, 2006).

While many authors advocate for the use of TEs in science instruction, there is little guidance available for teachers on how to apply them effectively and assist students in developing proper reasoning skills (Matthews, 1994; Reiner, 1998; Gilbert & Reiner, 2000); Hadzidaki, 2008; Galili, 2009). Current research has largely examined TEs as tools for improving conceptual knowledge and problem-solving, but their role in fostering reasoning and argumentation skills remains understudied. Additionally, there is limited exploration of how TEs might enhance the argumentation abilities of PPSTs, who are pivotal in teaching these skills to students. Bridging this gap is crucial to uncovering the full educational potential of TEs in teacher training and their broader impact on science education.

These studies underscore the potential benefits of incorporating TEs and argumentation into science education to enhance conceptual understanding and develop argumentation skills. However, they also highlight the need for more empirical research to fully understand and optimize these instructional strategies. The objective of this study is to investigate the impact of TE enriched worksheets on PSSTs conceptual understanding of mechanics, and their argumentation skills.

Method

Research model

This study adopted a pretest-posttest control group design with random assignment. In this experimental design, participants were randomly divided into experimental and control groups (Büyüköztürk et al., 2018). The experimental group received the physics laboratory through the worksheets enriched with TEs whereas the control group received the regular physics laboratory. Conceptual understanding form were administered to both groups before and after the treatment of the worksheets.

Participants

The study was conducted with 35 first-year PSSTs enrolled in of the Physics 1 Laboratory course at education faculty of one of the state universities located in western part of Turkey. PSSTs were randomly assigned to the experimental (17 PSSTs) and control (18 PSSTs) groups. All participants fully engaged in the activities and data collection process.

Treatment

The PSSTs in both groups enrolled the study for total of 10 hours in the physics laboratory for 4 weeks. Experimental and control groups participated the study in a different weekly schedule. The PSSTs in both groups worked individually. The researchers conducted the study by guiding the PSSTs in the framework of the worksheets. Both experimental and control groups received worksheets addressing core mechanics topics, in line with the objectives of the the central teacher training program of Council of Higher Education (YOK, 2018). Worksheets for the experimental group incorporated TEs with the topics covered, asking PSSTs hypothesize outcomes and explain their reasoning based on pre-existing knowledge about the TEs. PSSTs then conducted a regular (traditional) experiment on the related topic, collected data, and reflected on the hypothesis, constructing arguments in a structured format. The topics (and related TEs) are outlined as follows: Newton's laws (Elevator TE), universal gravitation and friction force (Galileo's Frictionless Incline), momentum and impulse (Huygens' TE), interaction in two dimensions (Baseball TE). Each worksheet's TE was based on recognized physics TEs or adapted from exercises in Physics (Serway & Beichner, 2012) with expert review from faculty in science education.

Worksheets for the control group were the same as the experimental groups' except for TEs part. The PSSTs in the control group were also asked to hypothesize outcomes and explain their reasoning based on pre-existing knowledge for each topic without the TE component. After conducting the regular experiments on the same topics, they also collected data, and reflected on the hypothesis, constructing arguments in a structured format.

As mentioned above the design of the worksheets for both groups ended with argument construction. Since, the PSSTs in both groups had not any experience about argumentation and argument construction, they were given a presentation providing information on the argumentation model and argument components (data, claim, support, rebuttal, etc.) by the researchers. At the end of the worksheets; the PSSTs were guided by the questions such as "In the 'Let's Think and Experiment' section, what did you aim to prove through the experiments you conducted? (What is the hypothesis/claim you generated for your research question?)" , "What kind of data (information, evidence, thoughts, observations, etc.) do you have to convince others of the situation you are trying to prove?" , "What is the basis of the data you have? (What principles or scientific explanations form the basis of your data?)" , "Under what conditions is your claim valid?" , "What are the exceptional or unacceptable probabilities that persist even when the conditions supporting your claim are met?" .

Data collection tools

Before and after the treatment PSSTs' conceptual understanding in mechanics was evaluated by using "The Mechanics Open-Ended Conceptual Understanding Form (MOCUF) developed by researchers. This form assesses conceptual understanding of mechanics concepts relevant to the Physics I course in the central teacher training program of Council of Higher Education of Turkey (YOK, 2018). The form was based on literature on mechanics education (Bademci & Sari, 2014; Acar & Gürel, 2016) and the main course source book (Serway ve Beichner, 2012). The initial form was composed of 12 open ended questions which was reduced to 9 items after revising in the direction of the views of the 1 physics teachers and 2 science and physics educator teaching staff at faculty of education. After revising upon the feedback from experts, the form was piloted with 25 third-year students. The results of the pilot study indicated a Cronbach's alpha of 0.69, verifying its suitability for the study (Kılıç, 2016). The data obtained through MOCUF was evaluated by the holistic rubric developed by the researchers. The rubric provided scores from 0 point (there is no relevant answer) to 4 points (there is the correct answer to the problem and a complete explanation) gradually. The rubric was developed by considering the extent to which the question has been answered, the relationship between the answer and the explanation, and the adequacy of the explanations. After the examination of three experts in physics teaching and science teacher education, the rubric was given the final form and piloted for its functionality.

The other data collection tool was the worksheets – particularly the sections where the PSSTs were required to construct arguments involving claims, data, warrants, and rebuttals– revealing the argumentation skills of PSSTs. The data obtained through the worksheets were evaluated by the "Argument Skills Assessment Rubric" (ASAR) rubric adapted from Cho & Jonassen (2002). ASAR evaluates argumentation skills by analyzing responses in terms of key components: claim, evidence/data, warrant, backing, qualifier and

rebuttal. Scores range from 0 to 6 for each component, with a maximum score of 36, enabling a thorough assessment of argumentation skills.

Table 1. Criterias and Quality Scores of Argument Skills Assessment Rubric" (ASAR)

Criterias	Quality Scores			
	0	2	4	6
Claim	There is no claim, or there are vague assertions.	The claim is not clear.	The claim is not fully specified.	The claim is clear and fully specified
Evidence/Data	No evidence/data has been used.	The data is incorrect, weak, or incomplete.	The provided data is relevant to the topic but not fully explained.	The supporting data is complete, accurate, and relevant to the claim.
Warrant	No connection has been made between the data and the claim.	The connection between the data and the claim is incorrect, or the data is not appropriate or valid.	Data has been provided, but its relevance to the claim is unclear.	Data has been used to support the claim.
Backing	No reasons for the warrants have been provided.	The reasons for the warrants are incorrect and irrelevant.	The correct, valid, relevant, and specific reasons for the warrants have been provided, but the reasons are too general.	It presents the correct, valid, relevant, and specific reasons for the warrants.
Rebuttals	No situations where the claim is not valid have been specified.	Very few situations where the claim is not valid are specified, and the limitations are not explained.	The situations in which the claim is not valid are specified in a flawed way.	The situations in which the claim is not valid are fully and systematically specified.
Qualifier	No qualifier has been specified	The qualifiers are specified, but it is not relevant.	The qualifiers are specified by establishing the necessary connections, and it is partially valid.	The qualifiers are correctly and relevantly specified

The scores from 0 to 6 in the Table 1 reflect a progression from weak to strong argumentation across multiple criteria. A score of 0 indicates the absence of key components such as a clear claim, supporting data, a valid warrant, backing, rebuttals, or qualifiers. A score of 2 suggests an attempt to include these elements but with significant flaws, such as unclear claims, weak or incorrect evidence, and insufficient reasoning. At 4, the argument improves, with relevant data and reasoning provided, but some explanations remain incomplete or unclear. A score of 6 represents the highest level of quality, where all components—claim, evidence, warrant, backing, rebuttals, and qualifiers—are clearly articulated, fully justified, and systematically connected, demonstrating strong argumentative reasoning.

Data analysis

To assess conceptual understanding, the data obtained from MOCUF were quantified. The resulting quantitative data were analyzed using the SPSS 23 software package, and interpretations of the findings are presented in the relevant section. To analyze the data, first of all normality of the dataset was checked by the skewness and kurtosis values.

According to Tabachnick & Fidell (2013), skewness and kurtosis values between -1.5 and $+1.5$ indicate that the dataset follows a normal distribution. Moreover, a significance result (p -value) of the test being greater than $.05$ ($p > .05$) indicates that the dataset follows a normal distribution (Büyüköztürk, 2012). In the interpretation of the test analyses, the significance level was set at $.05$ (Büyüköztürk, 2012). After confirming that the data met the normality condition, an independent samples t -test was applied to compare the pre-test and post-test results of the experimental and control groups to determine significant differences. Additionally, a paired samples t -test was used to determine whether there were differences within the experimental and control groups themselves.

In the evaluation of argumentation skills, the qualitative statements corresponding to each sub-component of the argument schemes constructed by pre-service teachers in their worksheets were

assessed according to the criteria in the argumentation skills evaluation rubric and converted into quantitative data. The quantitative data on argumentation skills were evaluated in two rounds. The entire evaluation process, including the initial evaluation phase (1 month) and the final evaluation phase (1 month), was completed over a total of 2 months, with a 2-month gap between the initial and final evaluations. To minimize errors arising from the scorer and ensure scorer reliability, the average of the results from the first and second rounds was taken, and the consistency of the researcher's scoring process in both rounds was calculated as an agreement rate of 80% (Miles and Huberman, 1994; Patton, 2002). Additionally, to determine whether there was a significant difference between groups across multiple dependent variables in the evaluation of argumentation skills, MANOVA (Multivariate Analysis of Variance) was conducted. The normality analysis, Levene's test for homogeneity of variance, and Box's test, which are prerequisites for conducting MANOVA, were examined, and it was found that the assumptions of variance-covariance equality were met. MANOVA was applied to determine which experiment or experiments contributed to the differences in total scores. The analysis results of the study data are presented under headings in the following section.

Ethics committee approval

The research process, including data collection and treatment was conducted with ethical approval from the relevant ethics committee of Dokuz Eylül University Institute of Educational Sciences of DEU (Document number:17.12.2019.15).

Results

Results from MOCUF

Addressing the effect of worksheets enriched with TEs on PSSTs' conceptual understanding of mechanics, the responses of PSSTs to MOCUF were analyzed. To determine whether the data followed a normal distribution, skewness and kurtosis values were examined. Table 2 shows the descriptive statistics results of MOCUF.

Table 2. The Descriptive statistics of the scores from MOCUF

	Experimental Group		Control group	
	Pre-test	Post-test	Pre-test	Post-test
N	17	17	18	18
M	11.55	16.91	12.75	15.66
SD	2.87	3.46	2.08	3.00
Skewness	1.17	.98	.28	-.20
Kurtosis	1.23	.32	.73	-.69
p*	.078	.081	.867	.220

*significant at $p < .05$.

As seen from Table 2, the pre-test score average of the pre-service teachers in the control group was found to be 12.75, while the post-test score average was 15.66. For the pre-service teachers in the experimental group, the pre-test and post-test score averages were 11.55 and 16.91, also +1.5, use parametric tests (independent samples t-test) were used to compare the MOCUF score of the groups before and after the treatment.

When MOCUF pre-test scores of PSSTs in the experimental and control groups were compared, no significant difference was observed in their conceptual understanding of mechanics topics between the groups ($t(33) = 1.41, p < .05$) (Büyüköztürk, 2012). Moreover, administering t-test for the post-test scores of PSSTs for MOCUF, it was observed that the post-test scores of the PSSTs in the experimental group did not significantly differ from those of the PSSTs in the control group ($t(33) = 1.13, p < .05$).

Followingly, MOCUF pre-test and post-test scores of PSSTs in the control group were compared a significant difference was found between the scores, $t(17) = 5.26, p < .05$. Additionally, MOCUF pre-test and post-test score comparison of PSSTs in the experimental group were compared a significant difference was found between the scores, $t(16) = 7.73, p < .05$. It was also observed in Table 2 that the difference in pre and post tests MOCUF average scores for the PSSTs in the experimental group (5.36) was greater than that MOCUF average scores of PSSTs in the control group (2.91).

Results from ASAR

PSSTs were administered four experiments during the process, and they were asked to construct arguments related to the experiments in the final section of the worksheets. The total scores obtained by the PSSTs in the experimental and control groups at the end of the four experiments were calculated and compared. The skewness and kurtosis values were also calculated for the experimental group as .292 and -.725,

respectively, and for the control group, .784 and .293, respectively. These findings are considered to exhibit a normal distribution (Tabachnick & Fidell, 2013). In this context, an independent samples t-test, a parametric test, was applied to compare the total scores of the pre-service teachers' argumentation skills in the experimental and control groups (Table, 3).

Table 3. Comparison of Total Scores of ASAR Between the Experimental and Control Groups

Group	N	M	S	Sd	t	p*
Experimental	17	90.82	11.65	33	-2.33	.026
Control	18	81.94	10.87			

*significant at $p < .05$

When examining the data in Table 3, it is stated that there is a significant difference in the total scores of argumentation skills of the PSSTs in the experimental and control groups ($p = .026$). To determine which experiment or experiments contributed to the difference observed in the total scores, MANOVA was applied. Certain conditions must be met before running MANOVA (Şen, 2016). Firstly, the dataset should exhibit a normal distribution. This initial condition was met by confirming the skewness and kurtosis values of each experimental data being between -1.5 and +1.5, and the p-values being greater than .05, indicate that the data shows a normal distribution (Büyüköztürk, 2012). The relevant p-values and other descriptive data are presented in Table 4 ($p > .05$).

Table 4. The Descriptive Statistics of The Scores from ASAR Regarding The Experiments

	Experimental Group				Control Group			
	Exp.*1	Exp. 2	Exp. 3	Exp. 4	Exp. 1	Exp. 2	Exp. 3	Exp. 4
N	17	17	17	17	18	18	18	18
M	19.88	22.23	24.94	23.76	18.22	19.66	23.55	20.50
SD	4.58	3.78	3.52	4.52	2.86	5.05	3.97	4.63
Skewness	.002	-.716	-.294	.054	-.659	-.257	.902	.568
Kurtosis	-.777	-.475	-.245	-.995	.165	-.622	1.48	-.376
p**	.805	.056	.713	.139	.128	.803	.181	.107

Exp.*: Abbreviation for "Experiment"

**significant at $p > .05$.

Secondly, the equality of covariance matrices, one of the assumptions, is presented in Table 5 with the results of Box's M statistic for running MANOVA.

Table 5. Results of Box's Test of Equality of Covariance Matrices

Box's M	19.538
F	1.039
SD1	10
SD2	5163.441
p*	.076

*significant at $p > .05$.

Table 5 reveals, the assumption of equality of covariance across groups was tested to check the equality of inter-variable correlations by $p = .076 > .05$.

Thirdly, the variances should be homogeneous. The homogeneity of variances was examined using Levene's Test, and the results obtained are presented in Table 6.

Table 6. Results of Levene's Test of Equality of Variances

	F	SD1	SD2	p*
Experiment 1	3.312	1	33	.078
Experiment 2	1.264	1	33	.269
Experiment 3	.046	1	33	.832
Experiment 4	.010	1	33	.919

*significant $p > .05$.

The data in Table 6, shows that the p-values are greater than .05 for all experiments. Accordingly, it has been determined that the variances are homogeneous. Upon reviewing Tables 4, 5 and 6, it was confirmed that the prerequisites for MANOVA were met, and MANOVA was applied to determine which experiment or experiments contributed to the difference in total scores. The analysis results are presented in Table 7.

Table 7. MANOVA Results for ASAR

Independent Variable	Group	N	M	SD	df	F / t	p*
Experiment 1	Experimental	17	19.88	4.58	1-33	1.67	.205
	Control	18	18.22	2.86			
Experiment 2	Experimental	17	22.23	3.78	1-33	2.87	.10
	Control	18	19.66	5.05			

Experiment 3	Experimental	17	24.94	3.52	1-33	1.18	.284
	Control	18	23.55	3.97			
Experiment 4	Experimental	17	23.76	4.52	1-33	4.47	.043
	Control	18	20.50	4.63			

*significant $p < .05$.

Table 7 shows there is no significant difference between the scores of Experiment 1 ($F(1-33) = 1.67, p > .05$), Experiment 2 ($F(1-33) = 2.87, p > .05$), and Experiment 3 ($F(1-33) = 1.18, p > .05$) for the PSST in the experimental and control groups. However, it is reported that the data of Experiment 4 significantly differ between the experimental and control groups in terms of argumentation skills ($F(1-33) = 4.47, p < .05$).

Discussion and conclusion

The purpose of this study was to put forward the effects of worksheets enriched with TEs on both conceptual understanding and argumentation skills of PSSTs in physics laboratory. The results, related to conceptual understanding of PSSTs, reveal that there was no significant difference in the conceptual understanding in both experimental and control groups before the treatment ($t(33) = 1.41, p < .05$). This result indicates that the conceptual understanding of the two groups was equivalent at the beginning. At the end of the treatment, there was a significant increase in the post-test conceptual understanding scores compared to the pre-test scores within each group ($t(17) = 5.26, p < .05$ for control group. and $t(16) = 7.73, p < .05$ for experimental group). When the post-test conceptual understanding scores of the participants in both groups were compared at the end of the treatment, no significant difference was observed between the two groups ($t(33) = 1.13, p < .05$). The reason for the absence of a significant difference within the groups could be the natural result of the physics course they attended throughout the semester that is parallel to the laboratory sessions. This course may help the development of PSSTs conceptual understanding. Similarly, Acar & Gürel (2016), investigated the concept of inertia in mechanics, stated that students showed conceptual development from high school to university level, and this development paralleled the historical development of science.

Moreover, for participants in both groups involving to construct arguments after conducting the regular experiments on the same topics may also lead to this result. Because, from a cognitive perspective, formulating an argument is a conceptual process that can aid in developing an understanding of the concepts and fostering deeper comprehension of scientific concepts as many researchers figured out (Erduran et al. 2004; Katchevich et al., 2013). For example, a study by Aydeniz & Dogan (2016) investigated the impact of argumentation on pre-service science teachers' understanding of chemical equilibrium and found out that argumentation facilitated a more profound grasp of the subject matter. Similarly, research by Acar et al., (2015) examined how argumentation activities influenced prospective secondary science teachers' understanding of balanced forces. The study found that participants developed counterargument and rebuttal skills, which correlated with improvements in their conceptual knowledge. Furthermore, a study by Cigdemoglu et al., (2017) focused on the role of argumentation in teaching acids and bases. Over six weeks, 29 freshman pre-service science teachers engaged in argumentation practices. The findings revealed that these practices contributed to the participants' chemical literacy, particularly enhancing their knowledge and competencies related to acids and bases. These studies collectively suggest that incorporating argumentation into science teacher education can be an effective strategy for deepening pre-service teachers' conceptual understanding, thereby better preparing them for their future teaching roles.

However, when comparing the average scores of MOCUF of both groups, it was seen that the average scores of PSSTs in experimental group increased more than that of PSSTs' in control group (Table 2). This increase in the experimental group's scores can be attributed to the effect of the worksheets enriched with TEs. Similarly, in a study examining the effect of TEs on students' conceptual understanding, Çetinkaya (2019) demonstrated that activities about simple machines in mechanics based on TEs had a positive impact on students' conceptual understanding. This result supports the observed increase in the conceptual understanding scores of the PSSTs in the experimental group in our study. In another study, İnce et al. (2016) emphasized that developing TEs in physics contributes to the understanding of physics concepts as a method. Moreover, studies by Reiner et al., (2000) and Gendler (2007) support these findings, suggesting that TEs foster deeper comprehension by allowing learners to simulate theoretical scenarios. For example, they found candidates in the experimental group reported that they could more effectively grasp the idealized conditions of concepts such as frictionless surfaces and gravitational forces through mental simulations prior to laboratory work. In another study, Asikainen & Hirvonen, (2014) utilized the double-slit TE as a diagnostic tool to assess the understanding of both pre-service and in-service physics teachers. The findings revealed that many participants had partial understanding of fundamental physics concepts, highlighting the need for integrating TEs in teacher education to enhance conceptual comprehension. Kösem & Özdemir (2014) also investigated how TEs function as cognitive tools in solving conceptual physics

problems. These studies identified various structures and purposes of TEs, suggesting that they can be productive tools for students, including pre-service teachers, to advance their inquiry skills and conceptual understanding and supporting the obtained results, that physics TEs have a positive effect on conceptual understanding.

In the study, as mentioned before, questions aimed at developing argumentation skills were included in the final sections of the worksheets in both groups and evaluated by using ASAR. When the total ASAR scores obtained by the PSSTs in both groups were compared, a significant difference in favor of the experimental group was observed (Table 3). The results on the impact of TEs on PSSTs' Argumentation Skills based on analysis using the ASAR showed that the experimental group developed more robust argumentation skills.

In their study, Tüzün & Köseoğlu (2018) found that students in online groups could easily reconstruct TEs as new arguments and generate counterarguments during TE based argumentation processes. In another study, Tüysüz & Tüzün (2020) observed that students structured TEs as arguments in their reconstruction of astronomy-chemistry TEs, and this also contributed to their critical thinking skills, which include argumentation skills. TEs facilitated a structured approach to building scientific arguments, incorporating key elements of the Toulmin Model (claim, data, warrant, backing, qualifier and rebuttal). PSSTs were able to practice constructing arguments by hypothesizing outcomes in TEs and then supporting or refuting these hypotheses through empirical data collected in the lab. This dual-layered approach to reasoning and evidence gathering is consistent with research by Walton & Sartor (2013), highlighting the role of structured argumentation in strengthening critical thinking. The findings of another study conducted by Demiral & Çepni (2018) also support this result. The researchers examined the argumentation skills of science pre-service teachers on the topic of genetically modified organisms and concluded that the pre-service teachers' subject knowledge and critical thinking skills had a significant impact on their argumentation skills. Rengifo (2019) states that the understanding of very technical theories due to its high degree of abstraction. For example, he reports Newton's bucket experiment was built to demonstrate the absolute character of space and in this positive sense, TEs assume a pedagogical function. Therefore, more time may be required for PSSTs to gain this abstraction and critical thinking.

When analyzing which specific experiment(s) contributed to the difference in total ASAR scores between the two groups, it was determined that no significant difference in argumentation skills existed during the first three experiments; however, a statistically significant difference emerged between the groups in the final experiment (Table 7). This difference can be explained by the fact that the PSSTs might have faced more challenges in the first three experiments due to their lack of familiarity with the argumentation process and insufficient subject knowledge. Until the 4th experiment, they tried to construct arguments three times. This may suggest that it takes time for the effects of TEs incorporated into argument construction to become evident in argumentation skills. Similarly, literature on TEs also indicates that an increase in prospective science teachers' subject knowledge enhances their argumentation skills. (Sadler & Fowler, 2006; Acar, 2010; Khishfe, 2012; von Aufschnaiter et al., 2008). For example, Sadler & Fowler (2006), in a study conducted with 45 high school students, investigated how students used their subject knowledge during the argumentation process. They found that an increase in subject knowledge improved the quality of students' argumentation justifications. Comparing these results with our study, it is evident that PSSTs' insufficient subject knowledge in the first three experiments did not affect their argumentation skills, but after the third experiment, the positive effect of TEs on argumentation skills was observed alongside improvements in subject knowledge. If the number of experiments in the research process had been greater (or the treatment had lasted more than 10 weeks), this difference might have been more evident in experimental data. Chang & Lin (2024) investigated the effect of argumentation-based chemistry instruction on pre-service teachers' self-efficacy in argumentation over an 11-week intervention and reported that participants demonstrated a notable increase in their confidence to engage in argumentation within the context of chemistry education. Similarly, Erdogan & Kaya (2025) examined a 10-session instructional module aimed at enhancing pre-service science teachers' argumentation competencies. The findings indicated significant improvements in participants' written and oral argumentation skills, which they effectively transferred to real learning environments.

Conclusion and implications for science education

The findings of this study offer valuable insights into the potential of TEs as pedagogical tools in science teacher education. While both the experimental and control groups demonstrated similar improvements in conceptual understanding, the experimental group—engaged with TE-enriched worksheets—showed a significant enhancement in argumentation skills. This suggests that, while TEs may not immediately boost conceptual understanding in the short term, they effectively contribute to the development of higher order thinking skills, such as scientific reasoning and argument construction.

These results align with previous theoretical claims that TEs promote cognitive engagement, abstraction, and mental simulation (Nersessian, 1992; Reiner & Gilbert, 2000b). In the context of prospective science teacher training, these outcomes are particularly important, as future educators must possess not only scientific content knowledge but also the ability to facilitate reasoning-based discussions and promote argumentation in their classrooms (Erduran & Jimenez-Aleixandre, 2007). The observed improvement in argumentation suggests that incorporating TEs into laboratory settings can enhance pre-service teachers' ability to justify claims, evaluate evidence, and construct coherent arguments—skills central to 21st-century science education goals.

Furthermore, the study design—where participants worked individually with structured worksheets—demonstrates that even without collaborative dialogue, TEs encourage deeper reflection and critical analysis. However, since argumentation is inherently dialogic, future studies could benefit from integrating group-based activities to foster peer interaction and scientific discourse, further strengthening argumentation skills (Osborne et al., 2004).

The relatively short intervention period (10 hours over 4 weeks) may explain the absence of significant conceptual differences. It is possible that extending the intervention or providing more exposure to TE-based tasks could lead to more noticeable effects on conceptual understanding. This aligns with the notion that abstraction and the internalization of mental models require sustained engagement and iterative thinking (Matthews, 2014).

In conclusion, the results highlight the pedagogical value of TEs not only as cognitive tools but also as a means of supporting the development of critical scientific practices in teacher education. Their integration into laboratory activities presents a promising approach to enrich the scientific experiences of pre-service science teachers (PSSTs) and cultivate their capacity for evidence-based reasoning and scientific argumentation.

The findings also have important implications for science education, particularly in the context of pre-service physics teacher (PPST) training. While no significant difference in conceptual understanding was found between the experimental and control groups, the improvement in argumentation skills among the experimental group suggests that TEs are valuable pedagogical tools in physics laboratory settings. TE-enriched worksheets appear to provide cognitive scaffolding, supporting the development of reasoning and critical thinking.

However, the short duration of the intervention suggests that the full impact of TEs on both conceptual understanding and argumentation skills may not be immediately apparent. Developing these skills likely requires sustained exposure and practice over longer periods. Therefore, future programs should consider extending the duration of such interventions to foster deeper cognitive engagement and internalization of abstract reasoning processes.

Moreover, the individual work format used in this study may have limited opportunities for collaborative dialogue. Incorporating group work into laboratory activities could promote group discussion and peer-to-peer argumentation, which are known to support the development of scientific reasoning and discourse skills. Structuring laboratory sessions to include collaborative discussions around TEs could further enhance their effectiveness.

Overall, the results suggest that embedding TEs into laboratory worksheets—especially when combined with longer interventions and group-based learning environments—holds promise for fostering higher-order thinking and argumentation skills in pre-service science teachers. This approach could ultimately contribute to the preparation of educators who are better equipped to facilitate inquiry-based and reasoning-centered instruction in their future classrooms.

Declarations

Ethics statements

This study was approved by the ethics committee of Dokuz Eylül University Institute of Educational Sciences under approval number 17.12.2019.15. All participants were informed about the objectives of the study, the voluntary nature of their participation, and the confidentiality of their responses. Informed consent was obtained from all participants prior to data collection.

Data availability

The datasets analyzed during the current study are available from the corresponding author on request.

Competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Authors' contributions

This study is part of the first author's master's thesis, conducted under the supervision of the second author. Ozge Tagluktimur was responsible for the conceptualization, design, data collection, and initial drafting of the manuscript. Gul Unal Coban as the academic advisor, provided guidance throughout the research process, contributed to the methodological design, the interpretation of the results and revised the manuscript critically for important intellectual content. Both authors reviewed and approved the final version of the manuscript.

References

- Acar, H. (2003). *Fizik öğrencilerinin düşünce deneyleri ile düşünme süreçlerinin incelenmesi* [Investigation of physics students' thinking processes through thought experiments] [Unpublished doctoral thesis]. Marmara University, Institute of Educational Sciences.
- Acar, H., & Gürel, Z. (2016). Düşünce deneyleri yolu ile öğrencilerin eylemsizlik kavramının araştırılması [Investigation of students' concept of inertia through thought experiments]. *Boğaziçi University Journal of Education*, 33(2), 69–96.
- Acar, Ö., Patton, B. R., & White, A. L. (2015). Pre-service science teachers' development of argumentation skills and conceptual understanding of balanced forces. *Australian Journal of Teacher Education*, 40(9), 36–56. <https://ro.ecu.edu.au/ajte/vol40/iss9/8>
- Aktaş, M., & Doğan, S. (2018). Argumentation-based teaching and its impact on argumentation skills. *Science Education Review*, 21(2), 178–192.
- Altıok, O. (2017). TGA tekniğine dayalı laboratuvar etkinliklerinin Fen Bilgisi öğretmen adaylarının argüman oluşturma becerilerine etkisinin incelenmesi. Yüksek Lisans Tezi, Recep Tayyip Erdoğan Üniversitesi.
- Asikainen, M. A., & Hirvonen, P. E. (2014). Probing pre- and in-service physics teachers' knowledge using the double-slit thought experiment. *Science & Education*, 23(9), 1811–1833. <https://doi.org/10.1007/s11191-014-9710-1>
- Ateş, M. E. (2015). Bilimlerde düşünce deneyleri [Thought experiments in the sciences]. *Mediterranean Journal of Humanities*, 5(1), 125–138. <https://doi.org/10.13114/MJH.201511372>
- Aydeniz, M., & Doğan, A. (2016). Exploring the impact of argumentation on pre-service science teachers' understanding of chemical equilibrium. *Chemistry Education Research and Practice*, 17(3), 420–432. <https://doi.org/10.1039/C5RP00170F>
- Bademci, S., & Sarı, M. (2014). Thought experiment in solving physics problems: A study into candidate physics teachers. *Education and Science*, 39(175), 203–215.
- Berland, L. K., & McNeill, K. L. (2010). A learning progression for scientific argumentation: Understanding student work and designing supportive instructional contexts. *Science Education*, 94(5), 765–793. <https://doi.org/10.1002/sce.20402>
- Brendel, E. (2004). Intuition pumps and the proper use of thought experiments. *Dialectica*, 58(1), 89–108.
- Brown, J. R. (1991). *The laboratory of the mind: Thought experiments in the natural sciences*. Routledge.
- Brown, J. R. (2006). The promise and perils of thought experiments. *Interchange*, 37(1–2), 63–75.
- Bybee, R. (2000). Teaching science as inquiry. In J. Minstrell & E. H. van Zee (Eds.), *Inquiring into inquiry learning and teaching in science* (pp. 20–46). American Association for the Advancement of Science.
- Büyüköztürk, Ş. (2012). *Sosyal bilimler için veri analizi el kitabı* [Handbook of data analysis for social sciences] (17th ed.). Pegem Yayınları.
- Büyüköztürk, Ş., Kılıç Çakmak, E., Akgün, Ö., Karadeniz, Ş., & Demirel, F. (2018). *Bilimsel araştırma yöntemleri* [Scientific research methods]. Özyurt Matbacılık.
- Çetin, P. S. (2014). Explicit argumentation instruction to facilitate conceptual understanding and argumentation skills. *Research in Science & Technological Education*, 32(1), 1–20. <https://doi.org/10.1080/02635143.2013.850071>

- Chang, M., & Lin, H. (2024). The development of pre-service teachers' argumentation self-efficacy through argumentation-based chemistry instruction. *Chemistry Education Research and Practice*. <https://doi.org/10.1039/D3RP00337J>
- Çiğdemoğlu, G., Arslan, H. O., & Çam, A. (2017). The role of argumentation in teaching acids and bases to pre-service science teachers. *Chemistry Education Research and Practice*, 18(4), 612–630. <https://doi.org/10.1039/C6RP00167J>
- Çetinkaya, İ. (2019). Basit makineler ünitesi ile ilgili geliştirilen düşünce deneyi etkinliklerinin 8. sınıf öğrencilerinin kavramsal anlamalarına etkisi [The effect of thought experiment activities developed for the simple machines unit on 8th-grade students' conceptual understanding] [Master's thesis, Aksaray University, Institute of Natural Sciences].
- Cho, K. L., & Jonassen, D. H. (2002). The effects of argumentation scaffolds on argumentation and problem solving. *Educational Technology Research and Development*, 50(3), 5–22.
- Demiral, Ü., & Çepni, S. (2018). Examining argumentation skills of preservice science teachers in terms of their critical thinking and content knowledge levels: An example using GMOs. *Journal of Turkish Science Education*, 15(3), 128–151. <https://doi.org/10.36681/>
- Driver, R., Newton, P., & Osborne, J. (2000). Establishing the norms of scientific argumentation in classrooms. *Science Education*, 84(3), 287–312.
- Erdoğan, İ., & Kaya, E. (2025). Making argumentation-based learning and teaching happen: Exploring the development of pre-service science teachers' argumentation competencies. *Science & Education*. <https://doi.org/10.1007/s11191-024-00612-1>
- Erduran, S., Simon, S., & Osborne, J. (2004). TAPPING into argumentation: Developments in the application of Toulmin's argument pattern for studying science discourse. *International Journal of Science Education*, 88(6), 915–933.
- Erduran, S., & Jimenez-Aleixandre, M. P. (Eds.). (2008). *Argumentation in science education: Perspectives from classroom-based research*. Springer.
- Galili, I. (2009). Thought experiments: Determining their meaning. *Science & Education*, 18(1), 1–23.
- Gilbert, J. K., & Reiner, M. (2000). Thought experiments in science education: Potential and current realization. *International Journal of Science Education*, 22(3), 265–283. <https://doi.org/10.1080/095006900289741>
- Gilbert, J. K., & Boulter, C. J. (2000). *Developing models in science education*. Kluwer Academic Publishers.
- Hadzidaki, P. (2008). The Heisenberg microscope: A powerful instructional tool for promoting meta-cognitive and meta-scientific thinking on quantum mechanics and the 'nature of science'. *Science & Education*, 17(6), 613–639. <https://doi.org/10.1007/s11191-006-9057-3>
- Hofstein, A., & Lunetta, V. N. (2004). The laboratory in science education: Foundation for the 21st century. *Science Education*, 88(1), 28–54.
- Jiménez-Aleixandre, M. P., & Erduran, S. (2007). Argumentation in science education: An overview. In S. Erduran & M. P. Jiménez-Aleixandre (Eds.), *Argumentation in science education* (pp. 3–27). Springer. https://doi.org/10.1007/978-1-4020-6670-2_1
- Kaya, O. N., & Kılıç, Z. (2008). Etkin bir fen öğretimi için tartışmacı söylev [Argumentative discourse for effective science teaching]. *Ahi Evran Üniversitesi Kırşehir Eğitim Fakültesi Dergisi*, 9(3), 174–179.
- Karslı-Baydere, F., Ayas, A., & Çalık, M. (2020). Effects of a 5Es learning model on the conceptual understanding and science process skills of pre-service science teachers: The case of gases and gas laws. *Journal of the Serbian Chemical Society*, 85(4), 559–573. <https://doi.org/10.2298/JSC190329123D>
- Khishfe, R. (2012). Relationship between nature of science understandings and argumentation skills: A role for counterargument and contextual factors. *Journal of Research in Science Teaching*, 49(4), 489–514.
- Katchevich, D., Hofstein, A., & Mamlok-Naaman, R. (2013). Argumentation in the chemistry laboratory: Inquiry and confirmatory experiments. *Research in Science Education*, 43(1), 317–345.
- Kılıç, S. (2016). Cronbach'ın Alpha güvenirlik katsayısı [Cronbach's Alpha reliability coefficient]. *Journal of Mood Disorders*, 6(1), 47–48.
- Kösem, Ş. D., & Özdemir, Ö. F. (2014). The nature and role of thought experiments in solving conceptual physics problems. *Science & Education*, 23(4), 865–895. <https://doi.org/10.1007/s11191-013-9635-0>
- Kuhn, D. (1991). *The skills of argument*. Cambridge University Press.
- Kuhn, T. S. (1977). *The essential tension: Selected studies in scientific tradition and change*. University of Chicago Press.

- Larkin, D. (2017). Planning for the elicitation of students' ideas: A lesson study approach with preservice science teachers. *Journal of Science Teacher Education*, 28(5), 425–443. <https://doi.org/10.1080/1046560X.2017.1352410>
- Matthews, M. R. (2014). *Science teaching: The contribution of history and philosophy of science* (2nd ed.). Routledge.
- McDonald, C. V. (2014). Preservice primary teachers' written arguments in a socioscientific argumentation task. *Electronic Journal of Science Education*, 18(7), 1–22.
- McNeill, K. L., & Pimentel, D. S. (2010). Scientific discourse in three urban classrooms: The role of the teacher in engaging high school students in argumentation. *Science Education*, 94(2), 203–229.
- Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd ed.). Sage.
- Najami, N., Hugerat, M., Kabya, F., & Hofstein, A. (2020). The laboratory as a vehicle for enhancing argumentation among pre-service science teachers. *Science & Education*, 29(2), 377–393. <https://doi.org/10.1007/s11191-020-00107-9>
- National Research Council. (2012). *A framework for K–12 science education: Practices, crosscutting concepts, and core ideas*. National Academies Press. <https://doi.org/10.17226/13165>
- Nersessian, N. J. (1992). How do scientists think? Capturing the dynamics of conceptual change in science. In R. Giere (Ed.), *Cognitive models of science* (pp. 3–44). University of Minnesota Press.
- Norton, J. (1996). Are thought experiments just what you thought? *Canadian Journal of Philosophy*, 26(3), 333–366.
- 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.
- Özdem, Y., Ertepinar, H., Çakıroğlu, J., & Erduran, S. (2013). The nature of pre-service science teachers' argumentation in inquiry-oriented laboratory context. *International Journal of Science Education*, 35(15), 2559–2586. <https://doi.org/10.1080/09500693.2011.611835>
- Patton, M. Q. (2002). *Qualitative research and evaluation methods* (3rd ed.). Sage.
- Reiner, M. (1998). Thought experiments and collaborative learning in physics. *International Journal of Science Education*, 20(9), 1043–1058. <https://doi.org/10.1080/0950069980200903>
- Reiner, M., & Gilbert, J. K. (2000). Epistemological resources for thought experimentation in science learning. *International Journal of Science Education*, 22(5), 489–506.
- Sadler, T. D., & Zeidler, D. L. (2005). Patterns of informal reasoning in the context of socio-scientific decision making. *Journal of Research in Science Teaching*, 42(1), 112–138.
- Serway, R. A., & Beichner, R. J. (2012). *Fizik 1 [Physics 1]*. Palme Yayıncılık.
- Simon, S., Erduran, S., & Osborne, J. (2006). Learning to teach argumentation: Research and development in the science classroom. *International Journal of Science Education*, 28(2–3), 235–260.
- Sorensen, R. (1992). *Thought experiments*. Oxford University Press.
- Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). Pearson.
- Tüzün, Ü. N., & Köseoğlu, F. (2018). Bilim eğitiminde düşünce deneyleri temelli online argümantasyonla lise öğrencilerinin eleştirel düşünme becerilerinin geliştirilmesi [Developing high school students' critical thinking skills through thought experiment-based online argumentation in science education]. *Journal of the Turkish Chemical Society, Section C: Chemical Education*, 3(2), 77–98.
- Tüysüz, M., & Tüzün, H. (2020). The effect of astronomy-chemistry thought experiments based on argumentation on critical thinking skills of gifted students. *Erzincan University Journal of Faculty of Education*, 22(3), 818–836. <https://doi.org/10.17556/erziefd.668687>
- Unal, G. C. (2009). *The effects of model based science education on students' conceptual understanding, science process skills, understanding of scientific knowledge and its domain of existence: The sample of 7th grade unit of light* [Unpublished doctoral thesis]. Dokuz Eylül University.
- von Aufschnaiter, C., Erduran, S., Osborne, J., & Simon, S. (2008). Arguing to learn and learning to argue: Case studies of how students' argumentation relates to their scientific knowledge. *Journal of Research in Science Teaching*, 45(1), 101–131.
- Walton, D. N., & Sartor, G. (2013). Teleological justification of argumentation schemes. *Argumentation*, 27(2), 111–142.
- Yalçın, G. (2018). Sosyobilimsel biyoloji konularının fen bilgisi öğretmen adaylarının yazılı argümantasyon becerilerine etkisi [The effect of socioscientific biology topics on preservice science teachers' written argumentation skills] [Master's thesis, Bartın University]

Yükseköğretim Kurulu [YÖK]. (2018). *Fen bilgisi öğretmenliği lisans programı* [Science teaching undergraduate program].
https://www.yok.gov.tr/Documents/Kurumsal/egitim_ogretim_dairesi/Yeni-Ogretmen-Yetistirme-Lisans-Programlari/Fen_Bilgisi_Ogretmenligi_Lisans_Programi.pdf