Journal of Computer and Education Research
October 2023 Volume 11 Issue 22 (ISSN:2148-2896)
http://dergipark.org.tr/jcer

Research Article

The Influence of Concept Cartoon-Supported Issue of Demarcation-Based Scientific Argumentation on Argumentation Skills and Pseudoscientific Beliefs

Ümit DURUK *[1] Emine ÇAVUŞ [2] Abuzer AKGÜN [3]

[1] *Adıyaman University, Adıyaman, Turkey, uduruk86@gmail.com*

[2] *Şehit Fazıl Gürs Middle School, Adıyaman, Turkey, ecavus1980@hotmail.com*

[3] *Adıyaman University, Adıyaman, Turkey, aakgun@adiyaman.edu.tr*

* *Corresponding Author: uduruk86@gmail.com*

Article Info

Received: *05 July 2023*
Accepted: *04 October 2023*

Keywords: Pseudoscientific beliefs, issue of demarcation-based scientific argumentation, argumentation skills, retention, concept cartoon

10.18009/jcer.1323212

Publication Language: English

Abstract

Studies have found that argumentation may lessen students' pseudoscientific beliefs. However, few studies of argumentation have been handled in the context of pseudoscientific beliefs. The purpose of the present study was to examine the influence of concept cartoon-supported issue of demarcation-based scientific argumentation on middle school students' pseudoscientific beliefs. A pre-test/post-test quasi-experimental design including a control group was employed. Participants were 22 grade 7 Turkish female middle school students sampled from two classes of a single-sex school located at the edge of a southeastern town center in Turkey. When supported by the use of concept cartoons, issue of demarcation-based scientific argumentation significantly reduced the pseudoscientific beliefs with a large effect size and this decrease was retained even after 10 months. In addition, students' argumentation skills significantly improved. Instructors should offer metacognitive tools in pedagogical approaches along with argumentation to foster middle school students' argumentation skills and lessen their pseudoscientific beliefs.

To cite this article: Duruk, U., Çavuş, E. & Akgün, A. (2023). The influence of concept cartoon-supported issue of demarcation-based scientific argumentation on argumentation skills and pseudoscientific beliefs. *Journal of Computer and Education Research, 11* (22), 643-670. https://doi.org/10.18009/jcer.1323212

Introduction

Pseudoscientific claims (PCs) tend to be more common in cultures where critical thinking abilities/dispositions are not adequately developed due to poor quality of science education (Ede, 2000). Moreover, these unwarranted beliefs about pseudoscientific practices shaped in line with these claims are at an alarming rate among adolescents (Francis & Williams, 2009). In the school context, one of the pedagogical approaches known to be effective in reducing pseudoscientific beliefs is argumentation (e.g. Tsai, Lin, Shih & Wu, 2015). Argumentation engages students in more decidedly complex thoughts (Nussbaum &

Sinatra, 2003). In particular, argumentative discourse in the classroom enhances the development of rational thinking (Mercer, Dawes, Wegerif & Sams, 2004). In the pseudoscientific context, it can be said that one of the ways to develop rational thinking is discussions based on the issue of demarcation. In recent years, there have been many studies showing that practical studies on pseudoscientific claims and practices are carried out in this context (e.g. Es & Turgut, 2018; Kaplan, 2014; Turgut, Akcay & Irez, 2010). Discussions based on the issue of demarcation may scaffold the argumentation within the pseudoscience context. If science teaching manages these pedagogical goals, it will help students develop argumentation skills and reduced pseudoscientific beliefs (Chen, Wang, Lu, Lin & Hong, 2016).

As is well known, metacognition refers to regulation of one's own cognitive system and deals primarily with abstraction of new/existing cognitive structures (Dinsmore, Alexander & Loughlin, 2008). It may be stated that, among metacognitive strategies, concept cartoons will support the aforementioned scaffold in terms of deeper facilitation of learning, development of argumentation skills and achievement of their retention (Naylor & Keogh, 2013). Therefore, facilitating argumentation via a powerful tool that triggers higher-level thinking may precede higher-level outcomes (Schraw & Moshman, 1995). Metacognition, which carries this potential and is thought to be able to play a key role in the retention of understandings, skills and beliefs that are planned to be developed through argumentation, is one of the main components of argumentation in practice (Jimenez-Aleixandre & Erduran, 2007).

To sum, there are three key concepts guiding the design of this study: pseudoscientific beliefs, scientific argumentation, and metacognitive strategies. We reviewed the literature surrounding these particular concepts to better situate the study explained in the following sections in detail.

Theoretical Background

Pseudoscientific Beliefs

Pseudoscience is a set of ideas which claims to use scientific methods and procedures and thus be based on well-organised and durable information (Afonso & Gilbert, 2010) but establishes a faulty mechanism of reasoning by using individual anecdotes in the place of evidence (Carroll, 2005). PBs represent the version of confidence and dedication to this bunch of flawed and unwarranted beliefs put into action (Fasce & Picó, 2019). Students should be equipped with skills that could distinguish what is scientific from what is pseudoscientific from early ages. However, putting this suggestion into action involves the ongoing debate on how to distinguish science from pseudoscience. Naturally, initiatives regarding this boundary create the problem of issue of demarcation due to the fact that there is no clear-cut checklist that could guarantee this (Pigliucci & Boudry, 2013). In schools, it is difficult to demarcate science from pseudoscience pedagogically. It may be an exaggeration to suppose this problem will permanently prevent the use of pedagogical approaches. Using referents rather than the boundary itself may help solve this difficulty. Instead of separate investigation of related referents (e.g. testable, falsifiable), a holistic scientific system of thought may pave the way for further understanding. Science courses fail to develop logical thinking, which helps students separate scientific facts from beliefs and dogma (Blanke, Boudry & Pigliucci, 2016). So, students make decisions with the low-level rational thinking through science education, and therefore, as the boundary between PCs and scientific claims becomes even more blurred, they fail to distinguish these claims from those that are scientific (Ede, 2000).

One of the theories explaining the scientific thinking system is the cognitive-experiential self-theory (Epstein, 2003). According to this theory, in cases of uncertainty where the probability of tendency towards PBs to increase is higher, the method of processing two pieces of information that are independent but in interaction is used. Rational thinking is on a conscious level and based on logic. However, experiential thinking is preconscious and based on past experience. Naturally, while making a decision in such cases, students are expected to conduct reasoning over rational thinking. Otherwise, students would become more prone to be caught in the attractiveness of PCs. For conducting rational thinking, awareness should be raised in students on why we value the scientific view, seeing science as a distinctive and valuable way of constructing knowledge and focusing science teaching based on the evidence (Chen et al., 2016).

Science education should focus on the complex interactions between the dimensions of scientific literacy and different types of unwarranted beliefs to improve pedagogical strategies (Fasce & Picó, 2019). For the purpose of preventing the potential harms of pseudoscientific practices, the boundary between science and pseudoscience which is not clear-cut may be clarified further during the instruction at the school (Kaplan, 2014). The characteristics of PBs can also be determined by this way (Afonso & Gilbert, 2010). In this method, it may be useful to utilise scientific argumentation. Through argumentation, students learn to defend and justify their own ideas using warrants and look for why some claims are credible while others are not. So, scientific argumentation may be an effective instructional strategy to help students develop appropriate epistemic and social norms of science. Given the very nature of argumentation as being involved in the scientific culture in the context of developing epistemic criteria (Jiménez Aleixandre & Erduran, 2007), issue of demarcation-based scientific argumentation, which allows students to question the underlying assumptions of PBs via evidence-based discussions, was developed and utilised by the authors in the study.

Issue of Demarcation-Based Scientific Argumentation, Concept Cartoons and the Issue of Retention

Typically, a scientific argument refers to the justification of claims with reasons and/or evidence (Guilfoyle, Erduran & Park, 2020). For example, the claim 'day and night occur regularly' has been justified with the spinning earth (Erduran, Guilfoyle & Park, 2020). In parallel with the increase in the value given to science today, pseudoscientific arguments have also become widespread. For example, the beliefs that water dowsing is effective, crystals bring luck, aromatherapy cures illnesses, and communication with the dead through psychics are based on pseudoscientific arguments. These beliefs do not contain the materials, methods and standards involved in the process of producing scientific arguments (Carroll, 2005). For example, magnetic field strength decreases as distance increases, whereas the extra sensory perception (ESP) as a paranormal belief does not depend on distance. Scientific argumentation is considered as a social process that involves the critique, persuasion and revision of knowledge claims both through scientific and dialogical discourse (McNeill, Katsh-Singer, González-Howard & Loper, 2016). It is evident that scientific argumentation improves argumentation skills (Cetin, 2014). Few research have examined these skills in other contexts such as pseudoscience. Demarcation may strengthen PSI-based argumentation activities that aim to reduce PBs. The issue of demarcation philosophically poses a substantial foundation for discussion, and although there is no consensus yet, it provides science educators with crucial breakthroughs about science's process, output, and assessment (Turgut, 2009). The existing insolubility associated with the issue of demarcation and the probability that there will never be a solution do not necessarily require neglecting the inherent value of this context of debate (Turgut et al., 2010). Therefore, conducting instruction in a pseudoscientific context may provide opportunities for students to improve their argumentation skills. Despite the existence of studies on the topic of examining teachers' perceptions of science and of students, a significant basis of research focused on the issue of demarcation in a way to also involve philosophical debates has not been established, and there is a significant shortcoming in the ability to carry the fundamental narratives of philosophical initiatives into the field of science education (Turgut, 2009). In one of close studies, Cetinkaya, Turgut and Duru (2015) examined the effect of issue of demarcation-based discussions and found that eight grade students' beliefs in science could be developed through these discussions in the pseudoscientific context of iridology. In the study directly related to PBs and scientific argumentation, Tsai et al. (2015) examined the effect of scientific argumentation on the PBs. Some of the PBs of the students were significantly lower in both the post-test and the retention test compared to the control group. Cekbas and Ozel (2019) examined the effect of argumentation in the context of astronomy on the PBs of science teacher candidates. PBs decreased significantly in the treatment group following astronomy teaching without argumentation. Ultimately, in Arik and Akcay (2018) with seventh grade students, argumentation-based activities including cosmic science, reflexology, astrology, numerology, and bioenergy were supported by issue of demarcation-based discussions. It was observed that the knowledge and discussion skills of the students about the issue of demarcation increased through argumentation. Also, the students developed more criteria and made more comprehensive inquiries. Taken all together, it was concluded that PSIs such as astrology, iridology, and numerology were used with various scientific criteria during the issue of demarcation-based discussions and these were effective in reducing PBs. To sum, one may argue that debates held in pseudoscientific contexts have the origin of the issue of demarcation, and therefore, practices regarding this issue are prevalently carried out on this discussion of demarcation (Cetinkaya et al., 2015; Turgut et al., 2010) and the studies above has given rise to the idea of facilitating the instruction through "Issue of Demarcation-based Scientific Argumentation" (henceforth, IDSA treatment).

The IDSA treatment basically and operationally refers to scientific argumentation supported by issue of demarcation-based discussions and provides students pseudoscientific contexts to engage in these discussions during the course of the treatment. Theoretically, this treatment is based on two main theoretical frameworks. First, Jiménez Aleixandre and Erduran (2007) posit that argumentation is the activity of developing epistemic criteria within a social culture. Discussions structured on these criteria are of great importance in legitimizing scientific knowledge by means of the referents differentiated from those generally used in pseudoscientific contexts. Accordingly, the IDSA treatment was primarily designed to include student discussions during scientific argumentation based on the performance of basic skills (claim, data, and warrant) through the use of referents in the rational thinking including testable, falsible, analytic, logical, process-oriented and so on (Epstein, 2003). The main goal is to ensure that students become more familiar with these referents and understand the vital role of evidence in scientific argumentation. Second, the IDSA treatment was scaffolded with the notion of 'coordination' between claims and its supporting evidence. According to Kuhn (2005), a coordination between views and beliefs that both have epistemic origins and widespread in society (e.g. nature of science views, PBs) can be achieved (Afonso & Gilbert, 2010). During the treatment, students were urged to compare their current PBs with their newly gained scientific beliefs. As discussions proceed, students might be encouraged to evaluate the rationale behind their PBs and discover the related evidence is based on faulty reasoning (Osborne, Erduran & Simon, 2004).

Another issue to be emphasised here is the extent to which scientific argumentation will react to the use of metacognitive strategies that could play a significant role in increasing of argumentation skills in addition to reduction and retention of PBs. Nevertheless, the literature on this issue is rare. Metacognitive strategies may help overcome motivational problems, deeper understanding and retention of new content, skills and beliefs. In addition, there is also some evidence that student-student discussions in groups are more often related to subject matter and less to metacognitive strategies (Akerson, Morrison & McDuffie, 2006). To fill the gap, the IDSA treatment was supported by concept cartoons shown among metacognitive tools (Naylor & Keogh, 2013). It was aimed to reduce PBs throughout the IDSA treatment. Reducing these beliefs is more difficult than achieving the expected conceptual change at school level. Specifically, in order for the students to realize the faulty reasoning mechanisms underlying their PBs and to reconceptualise them within the scientific thinking, statements regarding scientific knowledge and PBs were presented through concept cartoons on a visual basis.

Considering studies making measurements of retention, it is seen that contexts such as subject matter and socioscientific issues (Khishfe, 2015) have been used. The number of studies examining retention of PBs is low (Tsai et al., 2015). As retention depends heavily upon under which context the study is carried out (Upadhyay & DeFranco, 2008), the present study is expected to provide further information about retention of these beliefs in a specific context and explore the effects of metacognitive strategies. The 10-month retention measurement in the study has been defined in the 'over few months' category. Since beliefs are generally more resistant to change, long-term retention interval and meaningful learning were taken into account to control the external threats to the measure of retention (Khishfe, 2015). Also, students with self-conscious ability and metacognitive awareness more retained their nature of science views (Akerson et al., 2006). Therefore, concept cartoons were used as metacognitive tool to make students aware of their learning process and scaffold meaningful learning.

Study Rationale

A related line of research in science education shows that studies based on argumentation have been prevalently carried out with students, teachers and prospective teachers from all age groups. Nevertheless, very few of these studies have been conducted with middle school students in the context of the issue of demarcation (Cetinkaya et al., 2015). Moreover, argumentation studies both was supported by usage of metacognitive strategies and examined on the level of single-sex middle schools have not been conducted thoroughly. So, little is known about single-sex school students' pseudoscientific beliefs and argumentation skills. Accordingly, there is much promise on treatments aimed at promoting argumentation skills and lessening pseudoscientific beliefs especially in middle school students. Thus, the resulting evidence could potentially help researchers studying the interaction between the concepts articulated in the rest of the study.

As suggested by Shah and Conchar (2009), single-sex education has the potential to increase student success in some contexts, which requires further exploration. What is more, this potential may be realised in settings where different instruction approaches are systematically planned and directly implemented, monitored and assessed. It may be stated that concept cartoons support deeper facilitation of learning and develop argumentation skills (Naylor & Keogh, 2013). Argumentation studies using concept cartoons and single-sex middle schools have not been undertaken thoroughly. The pseudoscientific beliefs (PBs) and argumentation skills of single-sex school students are relatively unknown. Accordingly, treatments that promote argumentation skills and reduce PBs among female students more prone to hold these beliefs show potential (Rice, 2003).

The present study adds a comparison group to delineate the influence of concept cartoons and focusing on female students in a single-sex middle school. Second, among others, the study aims to take the initiative holding the discussions carried out over pseudoscientific issues (PSIs) throughout the IDSA treatment including the basic components of argumentation and the development of these components. Lastly, the study is unique by investigating the long-term retention of PBs both by itself and when supported by concept cartoons. The key research questions guiding the study were:

(1) What is the influence, if any, of the IDSA treatment on PBs and their retention in control group?

(2) What is the influence, if any, compared to control group, of the IDSA treatment that is supported by concept cartoons on PBs and retention in treatment group?

(3) What is the influence, if any, of the IDSA treatment on argumentation skills?

Method

Research Model

In order to examine the influence of concept cartoon-supported IDSA treatment, we conducted a pre-test/post-test quasi-experimental design including a control group in the study. In this research design, participants are randomly allocated to a treatment and control group and scored on a test before and after the experimental manipulation (Stratton, 2019).

Participants

The study was carried out in a single-sex school located at the edge of a town centre. This small-size school was located in south-eastern Turkey. Grade 7 middle school students from two classes in the school, all of whom were female, participated in the study. They enrolled in a compulsory science course within the 2018-2019 academic year. Before the experimental manipulation, one of the classes was assigned as a control group (7-A class) and the other one was assigned as a treatment group (7-B class) randomly. The groups were coded as CG and TG, respectively. Both groups consisted of 11 students.

They were purposively recruited through homogeneous sampling that focuses on one particular subgroup (here girls) in which all the sample members are similar (Dugard & Todman, 1995). Studying with only female students was aimed at preventing situations arising from gender differences in PBs. Females are more inclined to beliefs in fortune-telling, while males are more inclined to paranormal beliefs (Metin, Cakiroglu & Leblebicioglu, 2020). Thus, possible differences were investigated by staying within the same sex. After deciding to study within the same gender, Rice's (2003) opinion on whether to study with females or males was taken as a basis and finally female students were studied. Ethical principles and rules were followed during the planning of the research, data collection, analysis, and reporting. The consent form was signed by the study participants.

Instructional Context

The treatment began with a 2-day orientation. On the first day, the students were introduced to Lizotte's argumentation quality framework through a presentation including Toulmin's basic argumentation components. The Toulmin Argument Model has basic and advanced components. According to McNeill and Pelletier (2012), for a qualified argumentation on the level of grades 6-8, the components of claim, data, and warrant as basic components are sufficient. Accordingly, the arguments generated by the students were analysed using an analytical assessment tool developed by Lizotte, Harris, McNeill, Marx and Krajcik (2003). The presentation was followed by a whole-class discussion about the components of an argument. On the second day, the students were given an informative case related to electrical circuits and asked to write down a claim, data and warrant.

One week after orientation, the treatment began. The treatment lasted 5 consecutive weeks within the science course that took place in the spring semester. In these weeks, PSIs of spoon bending (object deformation without physical force), graphology (personality analysis from handwriting), astrology (predicting personality traits and future based on planet movement), levitation (objects hovering without an apparent physical factor), and reflexology (massage application to hands and feet to heal organs) were used. The five PSIs employed in the treatment came from a factor analysis conducted by Tseng, Tsai, Hung, Liu and Huang (2008) about beliefs in fortune-telling, healing practices, and the paranormal.

The 1st stage of the main treatment consisted of three parts. The first 10 minutes of the first stage was designed as the preparation stage. Before each activity, the students were reminded of claims regarding the PSIs, and they were asked to share their knowledge about the issue. In the preliminary preparation part, it was ensured that they compared the information they collected as a result of their research to that collected by their peers. Afterwards, the second part of the first stage started. In this part, videos with an average length of 10 minutes including the speeches and images of pseudoscience experts or people claiming to realise a pseudoscientific claim from daytime programmes on Turkish TV channels were watched by the students. The students were asked then to take notes on the claims included in the videos. After this, the third part of the first stage started. This part also included a difference. The students were randomly divided into two groups as those with odd numbers in TG and those with even numbers in CG. In difference to CG, the argumentation forms in TG were prepared by using concept cartoons including competing theories (Osborne et al., 2004). While filling out the forms, the students in CG were asked to consider the videos, and those in TG were asked to consider both the videos and the concept cartoons. At the end of the first stage, the argumentation forms were collected from the students to be delivered back at the second stage.

At the second stage, both groups were divided into three groups each consisting of different students every week to discuss basic argumentation components. During the stage, students were made to defend their claims with evidence. The researcher went among the groups, directed the process without misguidance, and monitored in-group interactions because the students were young and lacked argumentation experience. After group discussions, students who wanted to change their forms could use a new color pencil. At the end of the second stage, argumentation forms were collected to determine their argumentation scores. Figure 1 summarizes the instructional context.

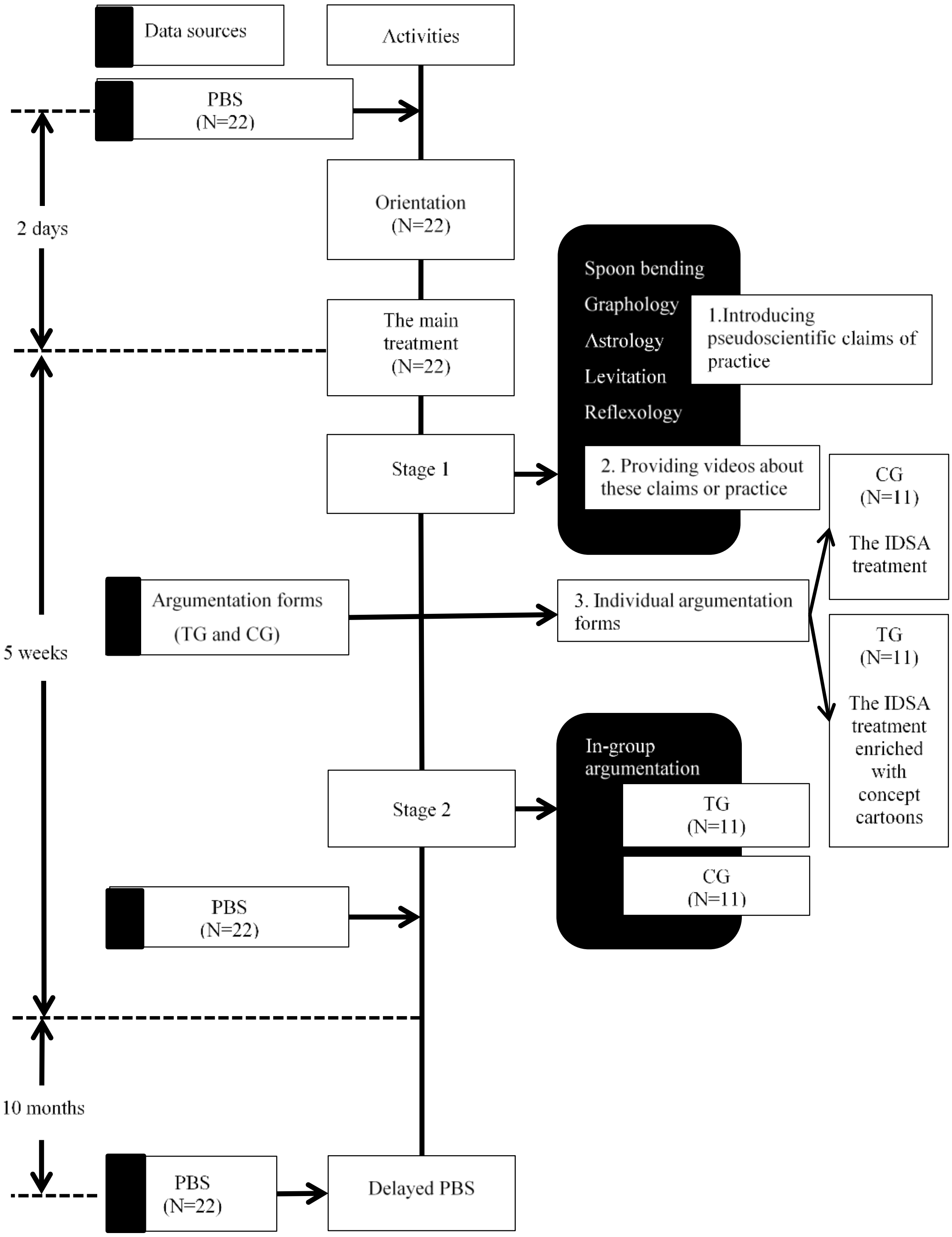

Figure 1. Study procedures and instructional context

Data Collection

Pseudoscientific Beliefs Scale (PBS) developed by Cetinkaya and Tasar (2018) was used to collect quantitative data. The minimum and maximum possible scores to be obtained in the scale are respectively 0 and 105. Higher scores indicate more PBs. The researchers found the Cronbach's alpha reliability coefficient of the entire scale as 0.84 while it was calculated 0.76 as reliable by pre-testing in the present study (Taber, 2018).

The five-point Likert-type scale consists of three dimensions as beliefs in fortune-telling, health practices and paranormal including 21 statements. Statements fall into three sub-dimensions: a) paranormal beliefs (1-9 items), b) fortune-telling (10-16 items), and c) health practices (17-21 items). The five scenarios used in the present study are directly associated with these three sub-dimensions take part in the scale. Spoon bending and levitation are paranormal beliefs, astrology and graphology are fortune-telling, and finally reflexology is a health practice.

To create the argumentation forms, the relevant literature was reviewed and pseudoscientific contexts were selected to help students write claims, data (evidence), and warrants. Although the contexts are the same, the forms given to TG included concept cartoons related to the relevant PCs. The concept cartoons included speech bubbles in a visual where four students stated their claims. One of the proposed claims included a scientific statement while the others included PCs. In the speech bubble containing the scientific statementin the concept cartoons, emphasis was made on the characteristics of scientific knowledge and why PCs cannot be scientific specified.

Data Analysis

The quantitative data were analysed using the SPSS 26 program. Because of normal distribution, independent-samples t-test was used to compare the scores of CG and TG, while paired-samples t-test was used to make comparisons among the pre-test, post-test and retention test results of the same groups.

The arguments were qualitatively analysed but quantitatively evaluated using an analytical assessment tool developed by Lizotte et al. (2003). High scores indicate higher levels of argumentation skills. The framework of analysis is given in Table 1.

Table 1. Argumentation skills quality framework (Lizotte et al., 2003)

Level	Claim	Evidence	Warrant
	Assertion or conclusion for a problem.	Data that support claim.	Argument that links evidence to claim.
0	Does not make a claim or make an inaccurate claim.	Does not provide evidence or only provides evidence that does not support the claim.	Does not provide reasoning or only provides reasoning that does not link evidence to the claim
1	Makes an accurate but incomplete claim.	Provides accurate but insufficient evidence to support the claim. May include some evidence that does support the claim.	Provides accurate and incomplete reasoning that links evidence to the claim. May include some reasoning that does not link evidence to the claim.
2	Makes an accurate and complete claim.	Provides accurate and sufficient evidence to support the claim.	Provides accurate and complete reasoning that links evidence to the claim.

All components were examined in terms of their scientific accuracy and completeness. Firstly, the process started with the analysis of the claims. The arguments of the students who argued that the claims based on the PBs in the concept cartoons were scientific or those who did not propose any scientific claim were scored zero, the arguments of those who stated that the PCs were not scientific but could not explain this completely were scored one, and the arguments of those who stated that PCs were not scientific and could completely explain this were scored two points. For example, the students who stated that a spoon cannot be bent with mental energy and said this can be achieved by magical talents rather than science received zero points for the claim component. Secondly, the evidence was analyzed. The arguments of students who supported their claim based on PBs with arbitrary evidence and those who agreed that their claims were not scientific but could not present accurate and sufficient evidence were scored zero, the arguments of those who agreed that their claims were not scientific but could provide insufficient evidence were scored one, and the arguments of those who agreed that their claims were pseudoscientific and could present accurate and sufficient evidence were scored two. Students who explained that people cannot fly and there is no evidence by saying scientists have not reached this conclusion scored one point in the evidence component. Finally, warrants were examined.

The arguments of students who said their pseudoscientific claim was scientific were scored zero, those who said it was not but failed to provide evidence were scored one, and those who accepted that their claims were not scientific and provided evidence were scored two. Students who said character analysis cannot be done by graphology, this claim must be supported by scientific evidence, and scientists cannot evaluate these statements based on observations or experiments scored two points in the warrant component.

The first and third authors analysed the data to guard against any bias that might result from the fact that the second author was the course instructor. To delve into reliability, three coders gathered before the analysis, and they read and compared the claims, evidence and warrants of all students. İlk analizde the Kappa statistic was calculated to determine the consistency among the coders for each component of the argument (i.e. claim, evidence and warrant). While the interrater reliability across the coders was relatively high for the claim (92%) with perfect agreement, it was relatively low on the evidence (72%) and warrant (66%) with substantial agreement (Landis & Koch, 1977). However, an iterative coding procedure was followed. Son analizde, these coding indices were compared and discussed in-depth to clarify ambiguities or discrepancies in the first analysis. Discrepancies were discussed until a consensus was reached. In addition, the SPSS 26 program was utilised to be able to statistically compare the qualitative data obtained from the argumentation forms. As the data did not show a normal distribution, Mann Whitney-U test for comparing the data of independent groups and Wilcoxon signed-rank test for comparing the data of dependent groups were used.

Results

Shapiro-Wilk test showed that PBS data were normally distributed ($p>.05$). Consequently, independent-samples t-test was used for the intergroup comparisons, while paired-samples t-test was used for the intragroup comparisons of CG and TG.

Table 2. Paired-samples t-test results on CG and TG

Group	Test	N	Mean	Std Deviation	Sd	t	p	Cohen's d
Control	Pretest	11	2.315	.5217	10	-1.721	.116	
	Posttest	11	2.050	.4182				
Treatment	Pretest	11	2.623	.4203	10	-2.334	.044*	1.146
	Posttest	11	2.009	.6299				

*p<.05

As seen in Table 2, while the posttest mean score of CG was lower than its pretest mean score, the difference was not significant (t=-1.1721, p>.05). The posttest mean score of TG was lower than its pretest mean score, and the difference was significant (t=-2.334, p<.05). Cohen's d value for this measurement as 1.146 refers to a high practical effect of the treatment. Consequently, concept cartoon-supprted IDSA treatment was highly effective in reducing PBs.

Table 3. The pretest, posttest and delayed test independent-samples t-test results for CG and TG

Test	Group	N	Mean	Std Deviation	Sd	t	p
Pretest	Control	11	2.315	.5217	10	1.529	.142
	Treatment	11	2.623	.4203			
Posttest	Control	11	2.050	.4182	10	.182	.858
	Treatment	11	2.009	.6299			
Delayed test	Control	11	2.190	.4363	10	-.354	.727
	Treatment	11	2.110	.6192			

*p<.05

As seen in Table 3, while the mean score of CG was numerically lower than that of TG, the difference between their pretest scores was not significant (t= 1.529, p>.05). This result showed that PBs of both groups were similar before the treatment. While the decrease in PBs of TG was numerically greater in comparison to that in CG, the difference between their posttest scores was not significant (t=.182, p>.05). In delayed test results, again, no significant difference was found between the mean scores of the groups (t=-.354, p> .05).

Table 4. Paired-samples t-test results on CG and TG

Group	Test	N	Mean	Std deviation	Sd	t	p
Control	Posttest	11	2.050	.4182	10	.827	.428
	Delayed Test	11	2.190	.4363			
Treatment	Posttest	11	2.009	.6299	10	.655	.527
	Delayed Test	11	2.110	.6192			

*p<.05

As seen in Table 4, although the mean delayed test score of CG was numerically higher than its mean posttest score, the difference was not significant (t=.827, p>.05). Similarly, although the mean delayed test score of TG was numerically higher than its mean posttest score, the difference was also not significant (t=.655, p>.05). This showed that concept cartoon-supported IDSA treatment achieved the retention of reduced PBs. Table 5 presents the scores of both groups in terms of their argumentation skills.

Table 5. Argumentation skill scores of CG and TG

		Control						Treatment					
		0		1		2		0		1		2	
		N	%	N	%	N	%	N	%	N	%	N	%
Spoon bending	Claim	6	54.5	4	36.4	1	9.1	11	100	0	0	0	0
	Data	8	72.7	2	18.2	1	9.1	11	100	0	0	0	0
	Warrant	8	72.7	2	18.2	1	9.1	11	100	0	0	0	0
Graphology	Claim	6	54.5	3	27.3	2	18.2	8	72.7	1	9.1	2	18.2
	Data	7	63.6	2	18.2	2	18.2	9	81.8	0	0	2	18.2
	Warrant	7	63.6	2	18.2	2	18.2	9	81.8	0	0	2	18.2
Astrology	Claim	4	36.4	5	45.4	2	18.2	8	72.7	2	18.2	1	9.1
	Data	6	54.5	3	27.3	2	18.2	10	90.9	0	0	1	9.1
	Warrant	7	63.6	3	27.3	1	9.1	10	90.9	0	0	1	9.1
Levitation	Claim	1	9.1	9	81.8	1	9.1	5	45.4	3	27.3	3	27.3
	Data	3	27.3	7	63.6	1	9.1	6	54.5	3	27.3	2	18.2
	Warrant	5	45.4	6	54.5	0	0	6	54.5	3	27.3	2	18.2
Reflexology	Claim	2	18.2	4	36.4	5	45.4	5	45.4	2	18.2	4	36.4
	Data	3	27.3	4	36.4	4	36.4	5	45.4	3	27.3	3	27.3
	Warrant	4	36.4	6	54.5	1	9.1	7	63.6	1	9.1	3	27.3

It is seen in Table 5 that the students in CG produced better arguments than the students in TG in the first week of the treatment. While TG produced weak arguments in the first two weeks of the treatment, their argument skill scores started to increase towards the last weeks. Figure 2 presents the total argumentation scores of CG.

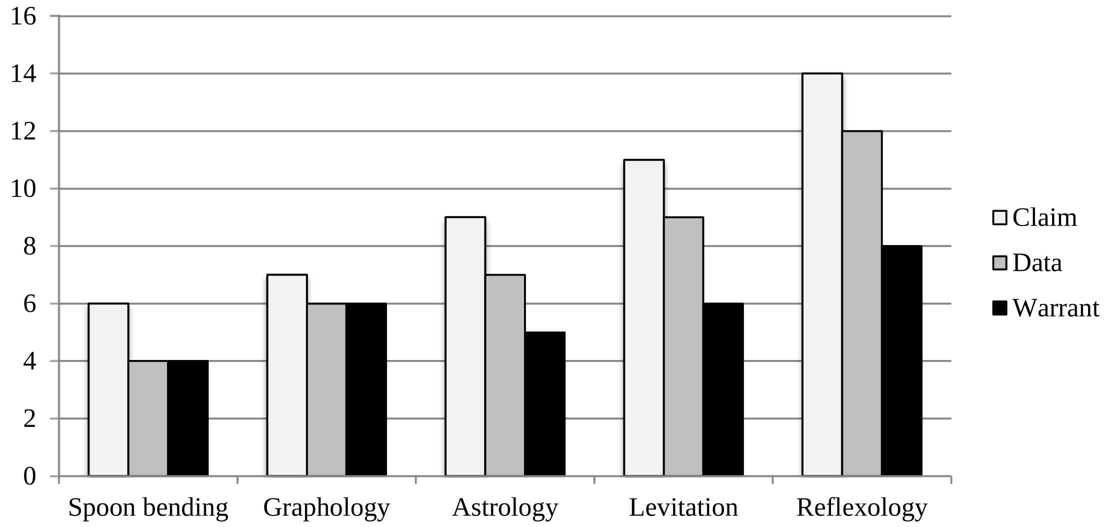

Figure 2. Graph of the total scores of CG regarding basic argumentation components

According to Figure 2, the scores of the basic components had a tendency to increase throughout the IDSA treatment. The highest increase was in the claim component, followed by data and warrant, respectively. However, the increasing trend in the claim component was very small, and the warrant component showed a decrease after the first two weeks where astrology was discussed.

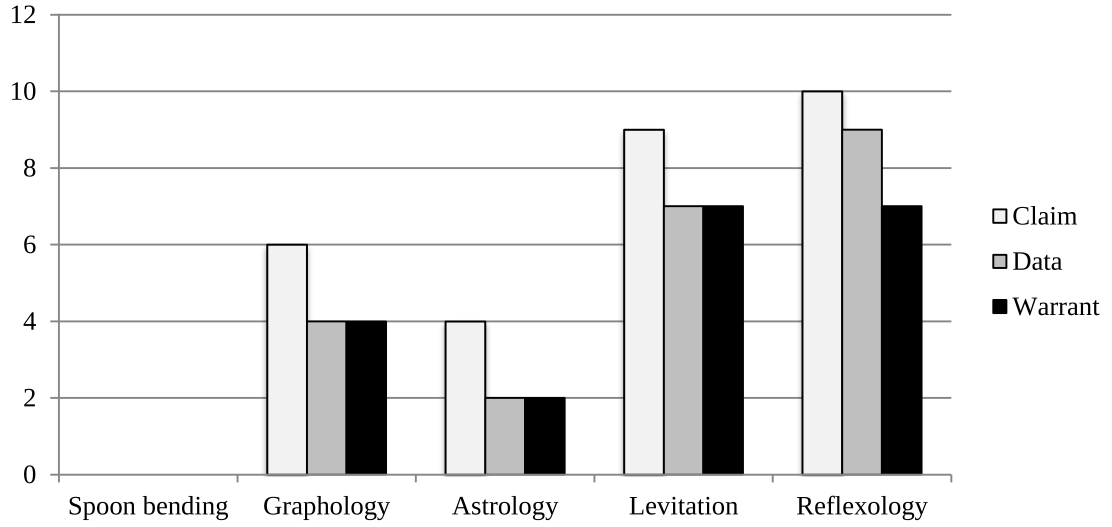

Figure 3. Graph of the total scores of TG regarding basic argumentation components

According to Figure 3, the scores of the basic components had a tendency to increase throughout concept cartoon-supported IDSA treatment. The highest increase was in the claim component, followed by data and warrant, respectively. However, the increasing trend in the basic argumentation components was not present. Similarly, the components showed a decrease after the first two weeks where astrology was discussed.

Shapiro-Wilk test did not show normal distribution ($p<.05$). Consequently, Mann-Whitney U test was used for the intergroup comparisons, while Wilcoxon signed-rank test was used for the intragroup comparisons of CG and TG.

Table 6. Wilcoxon test results on the first week and last week total argumentation scores of CG and TG*

Group	Argumentation components	Spoon bending-Reflexology	N	Mean Rank	Sum of Ranks	Z	p
Control	Claim	Negative Rank	2	4.50	9.00	-1.999	.046
		Positive Rank	8	5.75	46.00		
		Equal	1				
	Data	Negative Rank	1	6.50	6.50	-1.651	.099
		Positive Rank	7	4.21	29.50		
		Equal	3				
	Warrant	Negative Rank	1	5.50	5.50	-1.081	.279
		Positive Rank	5	3.10	15.50		
		Equal	5				
	Total	Negative Rank	2	5.00	10.00	-1.788	.074
		Positive Rank	8	5.63	45.00		
		Equal	1				
Treatment	Claim	Negative Rank	0	.00	.00	-2.271	.023
		Positive Rank	6	3.50	21.00		
		Equal	5				
	Data	Negative Rank	0	.00	.00	-2.251	.024
		Positive Rank	6	3.50	21.00		
		Equal	5				
	Warrant	Negative Rank	0	.00	.00	-2.251	.024
		Positive Rank	6	3.50	21.00		
		Equal	5				
	Total	Negative Rank	0	.00	.00	-2.232	.026
		Positive Rank	6	3.50	21.00		
		Equal	5				

**Based on negative ranks.*

Looking at Table 6, the last week scores of CG were numerically higher in comparison to its first week scores, and while this difference was significant for the claim component (z=-1.999, p<.05), it was not significant for the data component (z=-1.651, p>.05), warrant component (-1.081, p>.05) and the combination of the three components (z=-1,788, p>.05). While the last week scores of TG were higher in comparison to its first week scores, the difference was significant for the claim component (z=-2.271, p<.05), data component (z=-2.251, p<.05), warrant component (z=-2.251, p<.05) and the combination of the three components (z=-2.232, p<.05). In this context, the IDSA treatment supported with concept cartoons was effective in the increase of the argumentation skills. There was a high level of improvement in TG, whereas there was no significant difference between the groups in terms of their last week scores.

Table 7. Mann Whitney U test results of the first week and last week argumentation scores of CG and TG

			N	Mean Rank	Sum of Ranks	U	p
Spoon bending	Claim	Control	11	14.00	154.00	33.000	.013
		Treatment	11	9.00	99.00		
	Data	Control	11	13.00	143.00	44.000	.069
		Treatment	11	10.00	110.00		
	Warrant	Control	11	13.00	143.00	44.000	.069
		Treatment	11	10.00	110.00		
	Total	Control	11	14.00	154.00	33.000	.014
		Treatment	11	9.00	99.00		
Reflexology	Claim	Control	11	12.73	140.00	47.000	.345
		Treatment	11	10.27	113.00		
	Data	Control	11	12.50	137.50	49.500	.444
		Treatment	11	10.50	115.50		
	Warrant	Control	11	11.32	124.50	58.500	.887
		Treatment	11	11.68	128.50		
	Total	Control	11	12.18	134.00	53.000	.614
		Treatment	11	10.82	119.00		

Considering Table 7, when both group scores were compared for the first week of the treatment, there was a significant difference in the claim component (U=33.000, p<.05) and the combination of the three components (U=33.000, p<.05). On the other hand, there was no significant difference in the data component (U=44.000, p>.05) and the warrant component (U=44.000, p>.05). When the groups were compared at the last week, no significant difference was found in the claim component (U=47.000, p>.05), data component (U=49.500, p>.05), warrant component (U=58.500, p>.05) and the combination of the three components

(U=53.000, p>.05).

Discussion & Conclusion

Data analysis showed that IDSA treatment reduced PBs of CG (RQ1). In contrast to Cetinkaya (2017), the results of the present study are in keeping with the few research that found dramatically reduced PBs and retained reduced levels after incorporating scientific argumentation in pseudoscientific contexts (Tsai et al., 2015). Mixed results may have multiple reasons. First, Cetinkaya (2017)'s insignificant results might be explained as their issue of demarcation-based discussions lacked argumentation. Tsai et al. (2015) who utilized basic argumentation components showed that online argumentation significantly lowered students' PBs and this reduction was retained. Asking students to coordinate PCs and scientific evidence may have reduced PBs. During group discussions, instructional scaffolding (rational thinking referents) lead to the use of corroborative evidence to support assertions, which may inspire students to link PCs and their unsupportive evidence (McNeill & Pimentel, 2010). Claims and corroborating evidence may explain this. According to Kuhn (2005), coordination between newly learned scientific ideas and PBs may occur in the social learning context (Jiménez Aleixandre & Erduran, 2007). IDSA treatment created a social context where the epistemic underpinnings of science were discussed through the referents and claims, data, and warrants were scientifically examined. Third, a rubric with rising argumentation scores was applied (Lizotte et al., 2003). So, students might discuss about reasoning regarding PCs versus scientific claims, evidence, and warrants within each component (Khishfe, 2013). Students may have learned that each argumentation component's quality depends on its accuracy and completeness. This understanding may have prompted students to use rational reasoning more often when judging PCs (Chen et al., 2016; Epstein, 2003). Students who employed rational thinking references in the treatment may be able to demarcate science from pseudoscience. Students made connections between argumentation components they encountered during PSI discussions on science and pseudoscience. This could have helped them develop more qualified scientific justifications, which could have reduced their PBs (Tsai et al., 2015). Because argumentation is the process of producing warrants alongside evidence to examine claims (Kabapinar, 2020).

The second research question examines the effects of concept cartoons in support of TG. The significant changes between the pretest and posttest of TG were attributed to the

IDSA treatment including concept cartoons as a metacognitive tool. While PBs of TG declined, CG did not (RQ2). Results were consistent with the literature (Cekbas & Ozel, 2019; Tsai et al., 2015). The two studies cited above only supported RQ2 in terms of scientific arguments on PBs, not concept cartoons. We provide some reasons for the results. First, only PBs of TG decreased significantly. Concept cartoons can have cognitive and societal impacts. According to sociocultural constructivism, students should be able to participate in activities that turn their daily language into a scientific language. Students will endeavor to understand and interpret assertions and notions they express in their own words as scientific facts (Kabapinar, 2020). This significant decrease may be related to the concept cartoons used in small group conversations incorporated in the IDSA treatment's argumentative social environment, which helped students move from everyday language to scientific language (Osborne et al., 2004). Second, concept cartoons promote scientific awareness (Kabapinar, 2020). They allow students to understand the underlying reasons of their existing beliefs and participate in social conversation contexts ideal for this process (Keogh & Naylor, 1996). Thus, students in TG saw that their PCs were alternatives to scientific concepts, and they questioned the underlying causes of their own thoughts alongside other alternatives that developed in social dialogue. Cartoons may have helped TG students reason through more criteria. They may have challenged the experiential thinking system's of students to avoid falsifying evidence due to flawed reasoning (Carroll, 2005; Epstein, 2003). Third, concept cartoons are useful at triggering scientific argumentation (Keogh & Naylor 1996). They may have facilitated the coordination between claims and evidence in disproving alternative ideas by motivating students to gather corroborative evidence for the claims they defended (Kuhn, 2005). It may be thought that concept cartoons, utilized to create cognitive imbalance (Tsai et al., 2015) and stimulate conceptual change, were effective in reducing the students' PBs (Kabapinar, 2020).

The first two research questions also dealt with the retention of PBs. The reduction of PBs in both groups was retained after 10 months (RQ1 and RQ2). There are few long-term retention studies that have revealed similar results. Khishfe (2015) showed that many tenth-graders did not retain much of the change in their conceptual understanding after three months. Upadhyay and DeFranco (2008) discovered that students who received connected science teaching gained less environmental science knowledge in a short-term retention period but lost less after three months. Tsai et al. (2015) studied PBs retention after online argumentation. Even after a shorter retention interval, they identified fewer quantitative positive effects. Our results go beyond these mixed results. Future research should also examine the retention of argumentation skills in pseudoscientific contexts. Semb and Ellis (1994) found that the total loss of knowledge may not diminish as the retention interval rises. Our results corroborate this notion. These retained PBs need to be properly interpreted. First, it may be claimed that PBs of students arose from their prior knowledge rather than their formal science education at school because PSIs are not represented enough in Turkish middle school science textbooks (Duruk & Akgun, 2020) and were content-free in the IDSA treatment. Students may have had trouble connecting prior knowledge to conceptual change (McNeill & Pimentel, 2010). Despite this difficulties, both groups retained lower PBs. Retention is a product of meaningful learning based on continuing contact (Ausubel, 1962). Second, long-term retention depends on how well prior and new knowledge are related (Semb & Ellis, 1994). Accordingly, the high degree of retention in TG may be explained by the good organization of the connected pieces of knowledge created through the IDSA treatment and their longer-term storage in memory (Upadhyay & DeFranco, 2008). Third, domain exposure may be another factor (Khishfe, 2015). Continuous exposure to PCs may have helped students experience a cognitively more flexible conceptual change in time by helping them to structure their mental representations via meaningful learning (Kuhn, 1991). This cognitive flexibility may have helped students find suitable evidence for the claims through the concept cartoons. There is still a need for further studies examining both PBs and their retention and argumentation skills in content-embedded contexts, not generic ones.

The analysis showed that concept cartoon-supported IDSA treatment improved significantly the argumentation skills of TG in terms of all components. These results supported the claim that scientific reasoning can improve argumentation skills (Cetin, 2014). While CG's argumentation score increased insignificantly, TG's score did (Chen & She, 2012; Chen et al., 2016). While both groups received the IDSA treatment, only the claim component significantly increased in CG when concept cartoons were utilized (RQ3). This result is parallel with the studies posit that students put forward claims more easily than evidence and warrants (Chen & She, 2012; Osborne et al., 2004). Few research have focused on this difference's statistical significance (Chen & She, 2012; Chen et al., 2016). Hence, the results of this study are important in terms of revealing significant increase of argumentation scores in TG.

These results can be explained through scientific argumentation and related concepts. Argumentation activities help students apply knowledge and clarify their claims (McNeill et al., 2016). Through the discussions the boundary between science and pseudoscience carried out individually or in groups, students had the opportunity to practice evidence-based reasoning while they defending PCs against claims that did not have better evidence (Yeh & She, 2010). This treatment may have helped students develop a more rational thinking process (Epstein, 2003). Students lacking a basic epistemology of scientific practices do not consider their claims are untenable, thus they focus on their own claims and do not attempt to persuade others (Sandoval & Millwood, 2007). This may affect negatively how they use warrants to link evidence to a claim (Kuhn, 2005). When CG students realized their claims during group discussions were inaccurate, they may have been disturbed and avoided speaking up. With concept cartoons, the incorrect claims of the students in TG could be attributed to the cartoon character (Kabapinar, 2020). When students engage in argumentation, they experience multiple evidence-based perspectives (Khishfe, 2013). Concept cartoons may have helped present multiple perspectives together. By this way, students who hesitate to express the claims may develop cognitive flexibility, and they might be less resistant to change (Khishfe, 2013). Therefore, the provision of cognitive flexibility through concept cartoons (Kuhn, 1991) may have led to the significant increase of the evidence and warrant scores in TG (Maloney & Simon, 2006). Cognitively flexible students have less difficulty while producing more advanced argumentation skills (Khishfe, 2013). Future studies may profitably explore the effects of the IDSA treatment on argumentation skills by including students who are not directly received argumentation.

It was also evident that argumentation skills were transferable through different PSIs (RQ3). Except for astrology, PSIs in the study were relatively unfamiliar for the students (Sadler & Zeidler, 2004). Khishfe (2013) argued that skill transfer in unfamiliar contexts is harder. However, the students transferred their argumentation skills throughout these unfamiliar pseudoscientific contexts of the present study. This result was important in terms of demonstrating the possibility of transferring argumentation skills across PSIs. This result also resonates with the notion of McNeill et al. (2006) that claim-evidence-warrant structure of scientific argumentation supported the students' construction of scientific explanations. Argumentation skills could be transferred by these structure-oriented scaffolds to across PSIs. In the study, skill transfer may have been easier with the presentation of claim, evidence and warrant components in practice during the orientation. Contrary to Khishfe (2013), we found that students' argumentation skills, which are generally increased from spoon-bending to reflexology, deviated from the general increasing trend in the case of astrology (see Figure 2 and Figure 3). This conflicting result may be due to the media's frequent coverage of gender-sensitive themes including astrology (Kaplan, 2014; Metin et al., 2020).

The study had some limitations. First, generalization of the results may be limited to the study's participants and by cultural context, as PBs depend on cultural background (Tsai et al., 2015). Second, gender-based differences exist in pseudoscience. However, this differences are not observed in health issues. Therefore, health issues are appropriate for both male and female students (Metin et al., 2020). Reflexology was only health-related issue in the study. Overall, the present results are from a single-sex class in a co-educational school. Adding a comparison group, using concept cartoons, incorporating basic argumentation skills components, and measuring long-term memory may strengthen PBs and argumentation skills of students in these schools. These results should encourage future research in single-sex courses and compare single-sex and mixed classes regarding PBs and arguments.

Acknowledgement

It has been confirmed by the researchers that the data used in this study dates back to before 2020.

Author Contribution Statement

Ümit DURUK: *Conceptualization, literature review, methodology, writing and translation.*

Emine ÇAVUŞ: *Conceptualization, literature review, methodology, implementation, data analysis and writing.*

Abuzer AKGÜN: *Conceptualization, literature review, methodology and writing.*

References

Afonso, A. S., & Gilbert, J. K. (2010). Pseudo-science: A meaningful context for assessing nature of science. *International Journal of Science Education*, *32*(3), 329-348. https://doi.org/10.1080/09500690903055758

Akerson, V. L., Morrison, J. A., & McDuffie, A. R. (2006). One course is not enough: Preservice elementary teachers' retention of improved views of nature of science. *Journal of Research in Science Teaching*, *43*(2), 194-213.

Arik, M., & Akcay, B. (2018). An effectiveness of engaging in argumentation on students' ability to demarcate science from pseudoscience. *Sakarya University Journal of Education*, *8*(1), 41-60. https://doi.org/10.19126/suje.338919 .

Ausubel, D. P. (1962). A subsumption theory of meaningful verbal learning and retention. *The Journal of General Psychology*, *66*(2), 213-224. https://doi.org/10.1080/00221309.1962.9711837 .

Blanke, S., Boudry, M., & Pigliucci, M. (2016). Why do irrational beliefs mimic science? The cultural evolution of pseudoscience. *Theoria*, *83*(1), 78-97.

Carroll, R. T. (2005). *Becoming a critical thinker. A guide for the new millennium*. Boston: Pearson Custom Publishing.

Cekbas, Y., & Ozel, M. (2019). The effect of astronomy activities regarding Walton argumentation on pseudoscience beliefs of science teacher candidates. *International Journal of Eurasia Social Sciences*, *10*(37), 981-994.

Cetin, 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 .

Cetinkaya, E. (2017). *The effect of argumentation based activities, designed in the context of demarcation problem, on 8th grade students' views about nature of science, their pseudoscientific beliefs and argumentation skills* [PhD diss.]. Gazi University.

Cetinkaya, E., & Tasar, M. F. (2018). Development of pseudoscience belief scale (PBS): Validity and reliability study. *Trakya Journal of Education*, *8*(3), 511-526. https://doi.org/10.24315/trkefd.336650.

Cetinkaya, E., Turgut, H., & Duru, M. K. (2015). The effect of the context of science, pseudoscience demarcation on the science perceptions of secondary school students: The case of iridology. *Education and Science*, *40*(181), 1-18. https://doi.org/10.15390/EB.2015.3127.

Chen, C. H., & She, H. C. (2012). The impact of recurrent on-line synchronous scientific argumentation on students' argumentation and conceptual change. *Journal of Educational Technology & Society*, *15*(1), 197-210.

Chen, H. T., Wang, H. H., Lu, Y. Y., Lin, H. S., & Hong, Z. R. (2016). Using a modified argument-driven inquiry to promote elementary school students' engagement in learning science and argumentation. *International Journal of Science Education*, *38*(2), 170-191. https://doi.org/10.1080/09500693.2015.1134849.

Creswell, J. W., & Clark, V. L. P. (2007). *Designing and conducting mixed methods research*. California: Sage Publications, Inc.

Dinsmore, D. L., Alexander, P. A., & Loughlin, S. M. (2008). Focusing the conceptual lens on metacognition, self-regulation, and self-regulated learning. *Educational Psychology Review*, *20*(4), 391-409. https://doi.org/10.1007/s10648-008-9083-6.

Dugard, P., & Todman, J. (1995). Analysis of pre-test-post-test control group designs in educational research. *Educational Psychology*, *15*(2), 181-198. https://doi.org/10.1080/0144341950150207.

Duruk, U., & Akgun, A. (2020). Representation of nature of science components across secondary school science textbooks. *Amasya Education Journal*, *9*(2), 196-229.

Ede, A. (2000). Has science education become an enemy of scientific rationality? *Skeptical Inquirer*, 24, 48–51.

Epstein, S. (2003). Cognitive-experiential self-theory of personality. In T. Millon, & M. J. Lerner (Eds.), *Handbook of psychology: Personality and Social Psychology*, 5, (pp. 159-184). John Wiley and Sons, Inc. https://doi.org/10.1002/0471264385.wei0507 .

Erduran, S., Guilfoyle, L., & Park, W. (2020). Science and religious education teachers' views of argumentation and its teaching. *Research in Science Education*, 1-19.

Es, H., & Turgut, H. (2018). Candidate classroom teachers' perceptions about being scientific in the context of pseudoscience. *Journal of Education in Science Environment and Health*, *4*(2), 142-154. https://doi.org/10.21891/jeseh.409497 .

Fasce, A., & Picó, A. (2019). Science as a vaccine. The relation between scientific literacy and unwarranted beliefs. *Science & Education*, *28*(1-2), 109-125.

Francis, L. J., & Williams, E. (2009). The dayton agenda contacting the spirits of the dead: Paranormal belief the teenage worldview. *Journal of Research on Christian Education*, 18, 20-35. https://doi.org/10.1080/10656210902751818 .

Guilfoyle, L., Erduran, S., & Park, W. (2020). An investigation into secondary teachers' views of argumentation in science and religious education. *Journal of Beliefs & Values*, 1-15.

Jiménez Aleixandre, M. P., & Erduran, S. (2007). Argumentation in science education: An overview. In *Argumentation in Science Education*, (pp. 3-27). Springer, Dordrecht.

Kabapinar, F. (2020). *Caricature and concepts cartoons in science education*. Ankara: Pegem A

Kaplan, A. O. (2014). Research on the pseudoscientific beliefs of preservice science teachers: A sample from astronomy-astrology. *Journal of Baltic Science Education*, *13*(3), 381–393.

Khishfe, R. (2013). Transfer of nature of science understandings into similar contexts: Promises and possibilities of an explicit reflective approach. *International Journal of Science Education*, *35*(17), 2928-2953. https://doi.org/10.1080/09500693.2012.672774 .

Khishfe, R. (2015). A look into students' retention of acquired nature of science understandings. *International Journal of Science Education*, *37*(10), 1639-1667.

Kuhn, D. (1991). *The skills of argument*. Cambridge: Cambridge University Press.

Kuhn, D. (2005). *Education for thinking*. London: Harvard University Press.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 159-174.

Lizotte, D. J., Harris, C. J., McNeill, K. L., Marx, R. W., & Krajcik, J. (2003). *Usable Assessments Aligned with Curriculum Materials: Measuring Explanation as a Scientific Way of Knowing*. Paper presented at the annual meeting of the American Educational Research Association, Chicago: IL.

Maloney, J., & Simon, S. (2006). Mapping children's discussions of evidence in science to assess collaboration and argumentation. *International Journal of Science Education*, *28*(15), 1817-1841. https://doi.org/10.1080/09500690600855419.

McNeill, K. L., Katsh-Singer, R., González-Howard, M., & Loper, S. (2016). Factors impacting teachers' argumentation instruction in their science classrooms. *International Journal of Science Education*, *38* (12), 2026-2046.

McNeill, K., & Pelletier, P. (2012, May). *Supporting claim, evidence, and reasoning across the grades and curriculum*. Paper presented at the meeting of Annual Meeting of The National Science Teachers Association, Indianapolis, IN.

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. https://doi.org/10.1002/sce.20364.

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. https://doi.org/10.1080/01411920410001689689.

Metin, D., Cakiroglu, J., & Leblebicioglu, G. (2020). Perceptions of eighth graders concerning the aim, effectiveness, and scientific basis of pseudoscience: The case of crystal healing. *Research in Science Education*, *50*(1), 175-202.

Naylor, S., & Keogh, B. (2013). Concept cartoons: What have we learnt? *Journal of Turkish Science Education*, *10*(1), 3-11.

Nussbaum, E. M., & Sinatra, G. M. (2003). Argument and conceptual engagement. *Contemporary Educational Psychology*, *28*(3), 384-395.

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. https://doi.org/10.1002/tea.20035.

Pigliucci, M., & Boudry, M. (2013). *Philosophy of pseudoscience: Reconsidering the demarcation problem*. Chicago: University of Chicago Press.

Rice, T. W. (2003). Believe it or not: Religious and other paranormal beliefs in the United States. *Journal for the Scientific Study of Religion*, *42*(1), 95-106.

Sadler, T. D., & Zeidler, D. L. (2004). The morality of socioscientific issues: Construal and resolution of genetic engineering dilemmas. *Science Education*, *88*(1), 4-27.

Sandoval, W. A., & Millwood, K. A. (2007). What can argumentation tell us about epistemology?" In *Argumentation in Science Education*, (pp. 71-88). Springer, Dordrecht.

Schraw, G., & Moshman, D. (1995). Metacognitive theories. *Educational Psychology Review*, *7*(4), 351-371. https://doi.org/10.1007/BF02212307.

Semb, G. B., & Ellis, J. A. (1994). Knowledge taught in school: What is remembered? *Review of Educational Research*, *64*(2), 253-286. https://doi.org/10.3102/00346543064002253.

Shah, S., & Conchar, C. (2009). Why single-sex schools? Discourses of culture/faith and achievement. *Cambridge Journal of Education*, *39*(2), 191-204. https://doi.org/10.1080/03057640902903722.

Stratton, S. J. (2019). Quasi-experimental design (pre-test and post-test studies) in prehospital and disaster research. *Prehospital and Disaster Medicine*, *34*(6), 573-574. https://doi.org/10.1017/S1049023X19005053.

Taber, K. S. (2018). The use of Cronbach's Alpha when developing and reporting research instruments in science education. *Research in Science Education*, *48*(6), 1273-1296. https://doi.org/10.1007/s11165-016-9602-2.

Tsai, C. Y., Lin, C. N., Shih, W. L., & Wu, P. L. (2015). The effect of online argumentation upon students' pseudoscientific beliefs. *Computers & Education*, 80, 187-197. https://doi.org/10.1016/j.compedu.2014.08.018.

Tseng, Y. C., Tsai, C. Y., Hung, J. F., Liu, C. J., & Huang, T. C. (2008). *Belief in Pseudoscience among Students of Technological University*. Paper presented at the 24th Symposium on Science Education, National Changhua Normal University, Changhua, Taiwan.

Turgut, H. (2009). Pre-service science teachers' perceptions about demarcation of science from pseudoscience. *Education and Science*, *34*(154), 50-68.

Turgut, H., Akcay, H., & Irez, S. (2010). The impact of the issue of demarcation on pre-service teachers' beliefs on the nature of science. *Educational Sciences: Theory and Practice*, *10*(4), 2653-2663.

Upadhyay, B., & DeFranco, C. (2008). Elementary students' retention of environmental science knowledge: Connected science instruction versus direct instruction. *Journal of Elementary Science Education*, *20*(2), 23-37. https://doi.org/10.1007/BF03173668.

Yeh, K. H., & She, H. C. (2010). On-line synchronous scientific argumentation learning: Nurturing students' argumentation ability and conceptual change in science context. *Computers & Education*, *55*(2), 586-602. https://doi.org/10.1016/j.compedu.2010.02.020.