



Argument mapping as a pre-writing activity: Does it promote writing skills of EFL learners?

Qing Liu¹ · Zhiying Zhong² · John C. Nesbit¹

Received: 5 June 2022 / Accepted: 28 July 2023 / Published online: 19 August 2023

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2023

Abstract

Argumentation is a complex intellectual skill essential for academic achievement in a range of scholarly disciplines. Argumentative writing demands effective coordination of language use, reasoning processes, and background knowledge about a given topic. It is a cognitively challenging activity, especially for students learning English as a Foreign Language (EFL). The Dialectical Map (DMap) is an open-source argument visualization tool that scaffolds argument analysis, evaluation, and construction. A quasi-experiment was conducted with 190 undergraduate EFL students to investigate the pedagogical value of DMapping as a pre-writing activity. The research also explored the effect of DMapping on student motivation in learning argumentation. We found the DMap functioned as an effective pre-writing tool for argumentative writing and promoted transfer of argumentation skills. DMapping facilitated more balanced arguments, and students reported positive experiences and attitudes toward learning argumentation with the DMap. This paper concludes with implications and future research directions for using the DMap as a cognitive tool to foster meaningful learning.

Keywords Argumentative writing · cognitive tool · English as a Foreign Language (EFL) · argument visualization

1 Introduction

The need for reasoned decision-making in our changing societies constantly exposes citizens to questions that require scientific ways of thinking and critical evaluation of information obtained from multiple sources (Jumariati et al., 2021; National Research Council, 1996; Sawyer, 2014). Argumentation skill plays a crucial role in this process (Andrews, 2015; Jiménez-Aleixandre & Puig, 2012; Noroozi, 2023).

✉ Qing Liu
qing_liu@sfu.ca

¹ Simon Fraser University, 8888 University Dr, Burnaby, BC V5A 1S6, Canada

² South China Agricultural University, Guangzhou, China

Whether performed individually (e.g., writing an argument essay) or by a group (e.g., a debate), argumentation is a cognitive process in which propositions are constructed to justify or critique a claim with the goal of increasing or decreasing the tenability of a contentious point of view (Nussbaum, 2011; van Eemeren et al., 1996). As a complex intellectual skill, argumentation is required in many academic, technical, commercial, legal, and social service vocations (Andrews, 2015; Iordanou & Rapanta, 2021; Jiménez-Aleixandre & Puig, 2012). It has been pointed out that “argumentative processes stand at the intersection of critical and creative thinking” (Glassner & Schwarz, 2007, p. 17). Students who understand the structure of arguments and fact-based reasoning are better able to think critically, participate effectively as reflective citizens in a democratic society, and resist the allure of charlatanism and demagoguery. Researchers have found that constructing arguments for or against a contested claim promotes scientific thinking (Nesbit et al., 2019; Jonassen & Kim, 2010) and more meaningful processing and organization of information (Dole & Sinatra, 1998; Hogan & Fisherkeller, 2000; Niu, 2016). Argumentation is regarded as a vehicle for advancing an individual’s critical and higher-order thinking skills (Jumariati et al., 2021; Kuhn & Crowell, 2011; Stapleton, 2011).

Argumentation has garnered substantial attention in education and curriculum reform initiatives throughout the world (Ghanbari & Salari, 2022; Luo et al., 2020). In 2006, the European Union (EU) highlighted the ability to engage in argument-based inquiry as one of the key competencies for lifelong learning. In 2012, the National Research Council in the United States incorporated ‘constructing and evaluating evidence-based arguments’ as a key transferable knowledge and skill in the twenty-first century. The Program for International Student Assessment (PISA 2018) recognized the importance of the ability to ‘interpret and evaluate data and evidence scientifically and evaluate whether the conclusions are warranted’, which underlies argumentative processes. Having recognized its increasingly significant role, the Ministry of Education in China gives considerable emphasis to the development of students’ critical thinking and argumentation skills as one of the desired outcomes of the ‘New Curriculum Reform’ to enhance students’ soft skills and competitiveness (Tan, 2020).

1.1 Challenges of acquiring argumentation skills

Although constructing arguments is often called for in professional work and is essential in a wide range of academic disciplines, many postsecondary learners struggle with argumentation (Ghanbari & Salari, 2022; Jonassen & Kim, 2010; Wolfe et al., 2009). A person’s ability to argue for his or her own beliefs and interests tends to develop throughout childhood, but the more sophisticated *dialectical* aspects of argumentation ability – understanding and appropriately responding to ideas forwarded by others – do not naturally develop with age. Rather, fluency with more advanced forms of argumentation progresses when learners participate in activities that exercise them (Kuhn et al., 1997; Ogan-Bekiroglu & Eskin, 2012).

At the classroom level, teachers and instructors may not be paying sufficient attention to nurturing students’ argumentation skills. The phenomenon has been observed

internationally in secondary education in countries such as Netherlands (Majidi et al., 2021; van Eemeren et al., 2015), Mexico (Mejía & Neve, 2018), and Japan (Yamamoto et al., 2014; Yamamoto & Kamiyama, 2018). Xie and So (2012) conducted interviews and classroom observations to explore three 4th year pre-service teachers' understanding, skill, and instructional practice of argumentation in secondary science classrooms in Mainland China. They found the teachers participating in their study demonstrated limited understanding of argumentation, poor performance on argument construction, and little knowledge about the role of argumentation in science education, leading to an absence of in-class activities fostering those skills. This finding was echoed in research by Zhao et al. (2023) which indicated that many teachers in China are not well prepared for teaching scientific argumentation to students.

Wingate (2012) contended that students come to university with poor understanding and skills in argumentation. In most undergraduate programs, argumentation is not explicitly taught or trained (Mitchell & Riddle, 2000), resulting in many students graduating with insufficient mastery of these crucial skills. Research has found that postsecondary students showed a strong tendency to construct one-sided arguments without acknowledging opposing views or rebutting counterarguments (e.g., Heitmann et al., 2014; Januin & Osman, 2021; Jumariati et al., 2021; Liu, 2020). This problem is pervasive in various subject areas across grades (Nesbit et al., 2019; Deng & Wang, 2017; Pei et al., 2017; Wolfe & Britt, 2008).

Many native English speakers demonstrate difficulties in argumentative writing. It may be more challenging for English as a Foreign Language (EFL) learners, especially those who lack experience in related genres of writing in their first language (Ghanbari & Salari, 2022; Jumariati et al., 2021; Liao & Liao, 2022) or struggle with the differences in the rhetorical systems between their first language and the target language (Bacha, 2010). Argumentative writing is a complex cognitive task that involves processing of genre and subject domain knowledge related to a given topic, effective organization of information as the basis of reasoning, and appropriate use of language to clearly justify and critique one's argument (Jumariati et al., 2021; Liao & Liao, 2022). When EFL learners are writing an argumentative essay on an unfamiliar topic that requires them to refer to and process information from supplementary material(s) to support argument construction, it might be more likely for them to experience cognitive overload or fail to productively engage in argumentation as one of the key learning outcomes (Liu, 2020).

1.2 Argument schema

As mentioned above, many learners find it difficult to acquire advanced argumentation skills (Deng & Wang, 2017; Jonassen & Kim, 2010; Wolfe et al., 2009), including pre-service and in-service teachers (Garcia Romano et al., 2021; Xie & So, 2012; Zhao et al., 2023). Insufficiently developed argument schemas may underly this learning challenge (Reznitskaya et al., 2001; Wolfe et al., 2009). Wolfe et al. (2009) maintained that argumentative writing “requires the engagement and coordination of several cognitive processes such as retrieving a schema and encoding information from sources” (p. 184). A schema represents a modifiable

mental framework comprising a network of interrelations among the constituent units of knowledge gained from past experiences (Rumelhart & Norman, 1976). An argument schema is a cognitive structure that represents the components of argumentation (e.g., claim, data, warrant, backing, rebuttal, and qualifier) and the relationships among them (Reznitskaya & Anderson, 2002). The components and relationships in argument schemas are thought to be similar to those found in commonly used models of argumentation such as the Toulmin model (Toulmin, 1958).

Argument schemas enable the composition of arguments and the detection, understanding, and recall of argument-related information (Anderson et al., 2001; Reznitskaya & Anderson, 2002). The ability of students to argue and think critically in any situation is dependent on the capability of their argument schemas to meet situational demands (Nussbaum, 2008; Walton, 2013).

1.3 Creating a constructive environment for learning argumentation

In this study, we considered two theoretical frameworks when creating an instructional environment to develop students' argument schemas. While the ICAP framework (Chi, 2009) describes the relative cognitive benefits of different modes of engagement with new ideas or practices, the theory of cognitive tools explains how students' use of a suitably designed educational technology can enhance their argumentation ability.

1.3.1 ICAP framework

Chi's ICAP framework established the premise that "as students become more engaged with the learning materials, from passive to active to constructive to interactive, their learning will increase" (Chi & Wylie, 2014, p. 219). Passive learning refers to receiving information from instructors or instructional materials (e.g., listening to an instructor-led presentation) without overtly doing anything related to intended learning outcomes. Active learning, which involves manipulation of new information (e.g., selecting key ideas), is presumed to excel passive engagement. The constructive mode of engagement in which learners generate ideas beyond what is presented to them and actively construct their own knowledge (e.g., self-explaining) is more advanced than active learning behaviours. Interactive activities that involve constructive dialoguing or co-construction of knowledge with learning partners (e.g., group discussion or collaborative mapping) hold the potential for inducing superior learning outcomes (Chi, 2009). Reflecting on the depth of processing framework by Craik and Lockhart (1972), Chi et al. (2018) related constructive and interactive modes to deeper thinking and associated active and passive modes with shallower processing of information to be learned. This research investigates a constructive approach to acquiring argumentation skills.

1.3.2 Cognitive tools

Cognitive tools are technologies that mediate instruction by engaging, enhancing, and extending a learner's cognitive abilities in the process of learning (Pea, 1985; Salomon et al., 1991). Pakdaman-Savoji et al. (2019) summarized three key attributes of cognitive tools, including (1) representing concepts or operations that learners find challenging to understand or acquire (Jonassen & Reeves, 1996); (2) enabling iterated interactions with the tool such that the represented knowledge will be internalized and used spontaneously (Salomon et al., 1991); and (3) sustaining distributed cognition by offloading operations to the tool such that cognitive resources are freed for self-regulation of learning and dealing with more complex aspects of the task (Lajoie, 1993). Cognitive tools are thought to be more helpful for learning a skill at earlier stages when more guidance is needed, for tasks that are difficult relative to the learner's ability level (Jonassen & Reeves, 1996; Salomon et al., 1991), or for learners exposed to complex tasks such as argumentation (Nesbit et al., 2019; Pakdaman-Savoji et al., 2019).

1.3.3 Argument visualization tools

Argument visualization tools (AVT) are cognitive tools designed to guide the development of argument schemas by scaffolding learners' analysis, evaluation, and construction of arguments (Nesbit et al., 2019). AVTs are intended to make the structure of an argument more explicit and salient than text representations of the same argument (Davies et al., 2021; Dwyer et al., 2012; van Gelder, 2015). According to Nussbaum and Schraw (2007), AVTs can assist students to “activate, strengthen, and refine their existing [argument] schemas or develop new ones” (p. 65).

Davies et al. (2021) traced the history of argument visualization back to Richard Whately's work in the nineteenth century. In his *Elements of Logic* (Whately, 1834/1826), the first form of an argument map presented in a tree-like structure was introduced, but it was not widely adopted due to the challenge of using it to represent complex arguments. In the twentieth century, building on the *Chart Method* proposed by Wigmore (1913, 1937), Beardsley (1950) further developed it into a standard model in which argument components and their relationships are marked by linked numbers to represent different forms of reasoning (i.e., serial, divergent, or convergent reasoning). This diagrammatic approach was subsequently refined by other scholars such as Thomas (1973) and Scriven (1976). The widely cited argument model devised by Toulmin (1958) also significantly advanced this field.

In the past few decades, the notion of argument visualization has gained widespread interest. A variety of computer-supported AVTs have been developed to more efficiently facilitate the process of mapping arguments. For instance, *Rationale*TM is a commercial AVT (van Gelder, 2007), in which boxes of different colours are used to represent various argumentation components (e.g., contention, reason, objection, and rebuttal) and related boxes are linked with inference cues such as

“because”, “but”, or “however” to aid understanding of the relational structure of arguments. Research has found that engagement in argument mapping scaffolded by *Rationale*TM promotes reading comprehension of argumentative essays (Chiang et al., 2016) and development of critical thinking and argumentation skills (Dwyer et al., 2012; Ngajie et al., 2020). Hege et al. (2017) developed an AVT that was integrated into a Virtual Patients system to foster acquisition of clinical reasoning skills in health care education. Clinical reasoning requires argumentation skills as it involves diagnosing and decision making based on evaluation and integration of information collected from a variety of sources. The AVT enabled students to transform a patient’s illness script into a visual map, with findings, differential diagnoses, tests, and therapy options as nodes. Students labeled the connections between related nodes with different weights/colours (e.g., red for “speaking against” and dark blue for “highly related”). If needed, they could access an expert’s map to self-evaluate the quality of their own. The final task was to write a summary of the information presented in the map in a textbox placed at the bottom of the panel. Although empirical evidence substantiating its pedagogical effectiveness is lacking, this visualization tool potentially directs students’ attention to not only the outcomes but also the process of clinical reasoning. A critical review of four AVTs including *Belvedere*, *Convince Me*, *Questmap*, and *Reason!Able* (van den Braak et al., 2008) concluded that these AVTs were suitable for improving students’ argumentation skills.

Nussbaum (2008) developed a printed argumentation vee diagram consisting of two columns for pro arguments and counterarguments, respectively. It allowed students to draw arrows between opposing arguments to prompt refutation and presented two critical questions (i.e., “Which side is stronger, and why?” and “Is there a compromise or creative solution?”) to promote argument-counterargument integration. The results of a four-week experiment showed that argumentation vee diagrams fostered integration of counterarguments into argument essays but the effect was not transferred to a new argument essay developed without the aid of a vee diagram.

Building on Nussbaum’s work, Niu (2016) designed a computer-based AVT called Dialectical Map (DMap) that prompts students to visually assemble the key components of an argument, including pro reason, con reason, evidence, warrant, and conclusion. Research evidence supports the use of DMap as an effective AVT to scaffold and enhance argumentation-based learning. For instance, learners randomly assigned to study a scientific text by constructing DMaps recalled more main ideas from the text, and, consistent with an argument schema hypothesis, they structured the recalled information in a more argumentative format (Nesbit et al., 2019; Niu, 2016). Liu (2020) found that learners who engaged in retrieval practice aided by the DMap were better able to apply what they learned to new problem contexts and present stronger arguments in response to a controversial social issue.

Despite an abundance of research, the mechanisms by which AVTs may and may not be effective are not firmly established. To this end, variations among different types of AVTs and the contexts in which AVTs enhance learning should be broadly researched. A closer examination of the literature reveals that the AVTs used in existing empirical studies are often formatted as node-link diagrams in which nodes contain claims or evidence and links represent the supporting or opposing

Fig. 1 Initial student interface of the DMap

relationships between them (e.g., Binks et al., 2022; Davies et al., 2021; Dwyer et al., 2013; Khartabil et al., 2021; Ngajie et al., 2020). However, there is little research investigating the effects of bilateral AVTs such as the DMap (see Fig. 1) that comprises two columns for pro arguments and counterarguments, respectively. Burch et al. (2020) reviewed work on empirical evaluation of graph visualizations and concluded that the visualization layout impacts learning performance. Eye-tracking data revealed that node-link diagrams become harder to navigate and interpret when the size or density of networks increases, resulting in extraneous cognitive load (Huang, 2007; Yoghourdjian et al., 2021). Abdelaal et al. (2022) discovered that bipartite graphs tend to be more effective for structural overview and comparison than node-link diagrams. AVTs that adopt a systematic spatial organization less likely induce cognitive overload or map shock (Adesope & Nesbit, 2013; Nesbit et al., 2019).

Given its unique characteristics (e.g., layout emphasizing dialectical contrasts, prompts for warrants, and ratings of argument strength), the DMap was selected for use in this study. As a bilateral AVT, it may resolve navigation and interpretation issues that often hinder learning and cognitive engagement in argument mapping, especially for EFL students who are less experienced or skilled in argumentation. The findings of this research can potentially contribute to systematic investigations (e.g., meta-analytical reviews) of how various design characteristics of AVTs

foster learning and argumentation and thereby inform further development of theory, design, and appropriate use of AVTs as cognitive tools in education.

2 Research purpose and questions

As discussed above, argumentative writing is a cognitively taxing activity, especially for EFL students, as it requires effective orchestration of reasoning processes, language use, and background knowledge of the topic to be discussed (Jumariati et al., 2021). The DMap is a cognitive tool designed to scaffold argument construction. Informed by the ICAP hypothesis, DMapping holds the potential for engaging students in a *constructive* learning environment conducive to acquisition of argumentation skills. This study aims to investigate the pedagogical value of using the DMap tool in Chinese EFL classrooms to help students acquire advanced skills in argumentative writing and increase student motivation in learning argumentation. In addition, it set out to explore the factors that impact students' generation of balanced arguments and the role of DMapping in promoting counterargumentation. More specifically, this study addresses the following research questions:

RQ1: What are the effects of DMapping, designed as a constructive activity, on learning and argumentation?

RQ2: What beliefs and attitudes do students have toward DMapping?

RQ3: How does DMapping promote counterargumentation?

3 Method

The DMap tool is an emerging learning technology and there is a lack of empirical evidence corroborating and elucidating its effectiveness in fostering learning and argumentation. A quasi-experiment was conducted, complemented by qualitative data collected through questionnaires, to better understand whether and how DMapping supports learning and argumentation.

3.1 Participants

This research was conducted in a public comprehensive university in south China. A letter of approval was issued by the research ethics board of that university on January 22, 2021.

The participants were enrolled in four sections of the College English I course set up for first-year non-English major students to develop general English skills, including listening, reading, grammar, writing, translation, and vocabulary development. The present study included a DMap group ($n=91$) and a comparison group ($n=110$). Each group consisted of two randomly assigned course sections.

Referring to a study (Al-Shaer, 2014) investigating the effect of visual mapping as a pre-writing activity on EFL students' argumentative writing ($d=0.77$), a priori analysis using G*Power suggested that 56 participants would be required to detect the same effect size with power=0.80 and at $p < 0.05$ in an experiment with two independent groups. Detecting a medium effect size ($d=0.50$) would require 128 participants.

The participants came from 17 undergraduate degree programs (e.g., Biological Sciences, Artificial Intelligence, and Tourism Management). This research involved 3 essential sessions: learning intervention, delayed posttest #1, and delayed posttest #2. A few students missed some sessions because of COVID-19 vaccination appointments. Those who did not participate in the in-class learning intervention session and/or both posttest sessions were removed from further analyses. Consequently, 105 participants in the comparison group and 85 participants in the DMap group were retained. Of the 190 participants, 117 were male and 73 were female, with ages ranging from 17 to 20 ($M=18.27$, $SD=0.50$). Because the students participating in this research were drawn from across the university and the participating EFL classes took place in dispersed locations, we judged there was little risk of treatment diffusion (Johnson & Christensen, 2008).

3.2 Materials and instruments

3.2.1 Quick placement test (QPT)

Considering the importance of English proficiency in predicting EFL learners' writing performance (Jo, 2021), we used a standardized test called the Quick Placement Test (QPT), published by Oxford University Press (2001), to assess participants' English proficiency. The paper-and-pen QPT (Version 1) contains 60 multiple-choice questions diagnosing students' vocabulary, grammar and reading skills, with 60 points in total. It took 30 min to complete the test. The reliability and validity of this standardized test has been investigated in previous studies (e.g., Westrick, 2005).

Per the standard of the Association of Language Testers in Europe, the QPT scores represent different levels of English proficiency: 0–17 for Beginner, 18–29 for Elementary, 30–39 for Lower Intermediate, 40–47 for Upper Intermediate, 48–54 for Advanced, and 55–60 for Very Advanced. In this study, the QPT score was included as a continuous variable.

3.2.2 Dialectical map (DMap)

The Dialectical Map (DMap) is an open-source online visualization tool that guides students through the process of argument construction. The DMap was initially developed by Niu (2016). Since then, its user interface has been redesigned and it has been developed into an LTI (Learning Tool Interoperability) component that can be integrated with Learning Management Systems like Canvas for easier access and management.

As shown in Fig. 1, the DMap is a cognitive tool that prompts students to construct well-founded, balanced arguments in response to a specific claim (e.g., wind power should be widely used) presented at the top of the map. More specifically, the DMap allows students to (a) present as many arguments and counterarguments as they wish, (b) provide evidence either supporting or opposing each reason, (c) construct a warrant to explain the connection between each piece of evidence and the reason it is linked to, (d) label if a piece of evidence supports or opposes the corresponding reason, (e) rate the strength of each argument based on its importance or relevance to the claim, and (f) link arguments and counterarguments directly opposing each other. After weighing all arguments and counterarguments presented in the DMap, students type an integrated conclusion in a textbox at the bottom of the map. This tool was used to scaffold the argument mapping activity the DMap group engaged in.

3.2.3 Instructor-led argumentation tutorial

The comparison group received an instructor-led presentation introducing the key components of good argumentation including claim, pro reason, con reason, supporting evidence, opposing evidence, warrant, and conclusion. The DMap group received a tutorial led by the same instructor presenting the key components of the DMap.

The tutorials for the two groups showed the same components of argumentation and examples. The only difference was that the tutorial set up for the DMap group included screenshots demonstrating how to use the DMap (e.g., how to add a pro reason and how to delete an argument). Each instructor-led tutorial lasted about 15 min.

3.2.4 Reading passage

All participants were assigned to read a text about wind power that was adapted from the reading passage used in Liu (2020) according to the participants' English proficiency and study time given to each group. The wind power text used in this study contained 1,019 words, 59 sentences, and 128 propositions that provided a rich and balanced source of information supporting and opposing the claim that wind power should be widely used.

The Flesch-Kincaid reading grade level was 10.00, indicating that the text was understandable for average 10th grade students. Since our participants were EFL students, we also consulted two instructors of the College English I course and piloted it with 5 first-year undergraduate classes to make sure they had no difficulty understanding the text.

3.2.5 Outcome achievement measures

There were three learning outcome measures conducted to address RQ1 — one vocabulary test and two argument essays. The vocabulary test was related to the learning activity the comparison group engaged in during the intervention

session. It included 15 words (e.g., exploit, ecosystem) used in the wind power text. Participants were required to explain the meaning of each word in English. Each group was given 15 min to work on the vocabulary test.

In the first outcome assessment session, the participants were also instructed to write an essay about wind power, which was associated with the argument map construction activity the DMap group engaged in during the intervention session. It asked participants to write an essay (no less than 150 words) to argue if wind energy should be widely used. All participants were told that they could refer to the wind power passage while writing the essay and include ideas relevant to their arguments but not mentioned in the wind power passage. Those in the DMap group had no access to the dialectical maps they created at the intervention phase even though it addressed the same topic.

In the second outcome assessment session, the participants were required to write an essay with at least 150 words that argued whether universities should increase use of computer technology for student learning. This topic was chosen to assess the application of acquired argumentative writing skills to a topic different from wind power. All students were assumed to have substantial experience or knowledge about the pros and cons of using computers in education such that they need not refer to additional information for argument construction. Since the participants were not provided with any reference materials relating to this essay topic and neither group engaged in any pre-writing activities, this test intended to measure transfer of argumentation skills.

3.2.6 Questionnaire

Questionnaires were used to address RQ2 and RQ3. All participants responded to three 5-point Likert items regarding argumentation components, for example, “Discussing counterarguments (i.e., the arguments of an opponent) strengthens my essay”. There were also two comment questions asking if they had included counterarguments in the wind power essay and in the computer technology essay, respectively, as well as the reasons for doing that.

The participants in the DMap group were invited to complete an additional set of 5-point Likert items related to learning experience with the DMap tool. Deng et al. (2021) validated the 28-item questionnaire with 192 undergraduate students and found that it could reliably measure student attitudes toward DMapping. In their report, the Kaiser-Myer-Olkin (KMO) measure of sampling adequacy was 0.93. Bartlett’s test of sphericity was significant, $\chi^2(378)=2924.15$, $p<0.001$. An Exploratory Factor Analysis identified four factors that characterized students’ attitudes toward DMapping – effects of DMapping on learning and motivation, self-efficacy for DMapping, preferences for DMapping as an instructional mode, and DMap usability (Deng et al., 2021). In our samples, Cronbach’s α for each of these factors ranged from 0.72 to 0.87, indicating acceptable internal consistency.

3.3 Scoring methods

The vocabulary test involved 15 definition questions, with 15 points in total. The answer to each question was assigned 0 or 1 point depending on if the meaning of the word had been accurately explained in English as instructed.

For the argument essays, the number of words included in each essay was recorded as essay length has been found to be a predictor of essay performance (Fleckenstein et al., 2020). In this study, we considered higher quality argumentation to be demonstrated in essays that provide more supporting arguments, counterarguments, and rebuttals. A propositional scoring method was used to code the argument essays (Bovair & Kieras, 1985). The essay was broken into a set of idea units, each reflecting distinct meaning. For instance, the sentence “wind power is clean and renewable” included two propositions: (1) wind power is clean, and (2) wind power is renewable. To further analyze the components of each essay, a coding scheme consisting of 3 variables was developed to categorize the propositions, as follows:

- **Pro Argument:** A pro argument is a reason or supporting evidence in defense of the position taken. A warrant that explained how the evidence supports the pro reason it corresponds to was counted as a pro argument. Each proposition coded as a pro argument was assigned 1 point.
- **Con Argument¹:** A con argument is a reason or evidence opposing the proposed claim. A warrant that explained how the evidence supports the con reason it corresponds to was counted as a con argument. Each proposition coded as a con argument was assigned 1 point.
- **Rebuttal:** A rebuttal points out possible flaws in a counterargument. Each proposition coded as a rebuttal received 1 point.

Below is an example illustrating how an excerpt from a computer technology essay was scored to produce 5 points:

I think universities should increase use of computer technology for student learning. Computer-based technologies can potentially make learning more fun [*pro argument*] and increase student motivation in learning [*pro argument*]. However, it might also create study distractions [*con argument*] that result in poor learning performance [*con argument*]. Instructors should deliberately design computer-assisted learning environments to avoid unnecessary distractions [*rebuttal*].

3.4 Coding reliability

Six coders participated in two training sessions (3 h each) on how to analyze argument essays based on the coding scheme developed for this study. Three were assigned to score the wind power essays and the other three were assigned to score the computer technology essays.

¹ In this paper, the terms *con argument* and *counterargument* are used interchangeably.

Table 1 Interrater reliability results

Essay coding group	Phase 1 ICC Result			Phase 2 ICC Result			Phase 3 ICC Result			
	Pro	Con	Rebuttal	Pro	Con	Rebuttal	Pro	Con	Rebuttal	
Wind Power	0.67	0.81	0.36	0.73	0.91	0.51	0.89	0.92	0.85	Independent Coding
Computer Technology	0.44	0.88	0.65	0.84	0.96	0.82	Independent Coding			

As shown in Table 1, a phased approach was applied to obtain sufficient interrater reliability measured as interclass correlation coefficients (ICC). In each phase, each coding group was assigned to co-code 30 randomly selected essays. After a coding group demonstrated sufficient reliability (ICC > 0.80), the remainder of the essays were independently coded.

3.5 Research design and procedure

This study was conducted in an authentic classroom setting. The College English I course ran over 16 weeks, involving 32 lectures. For each course section, two lectures (40 min each) were delivered on the same day every week, bridged by a 5-min break. This research involved three sessions that were conducted consecutively from Week 10 to Week 12, either in the first or second lecture of the day, depending on how it related to the lecture topic and schedule. The learning activities and the posttests used for this research were part of course requirements, and no extra incentives were provided for participation.

Figure 2 shows an outline of the research design and procedure. Session 1 was set up for research intervention activities. The comparison group received an instructor-led presentation of argumentation components that lasted about 15 min, and then each participant was given the wind power passage to read. They were instructed to carefully read the passage as they would engage in a learning activity related to it in the next class. A filler task was designed for the comparison group that would engage them with the semantics of the passage without giving them practice in argument construction. After reading the wind power text, the participants were asked to

Group	Session 1 Week 10 Lecture			Session 2 Week 11 Lecture		Session 3 Week 12 Lecture	
	Instructor-Led Argumentation Tutorial (15 mins)	Wind Power Text and DMapping (25 mins)	Wind Power Text and Vocabulary Exercise (25 mins)	Wind Power Essay (40 mins)	Vocabulary Test (15 mins)	Computer Technology Essay (40 mins)	Online Questionnaire (5-10 mins)
Comparison	Y		Y	Y	Y	Y	Y
DMap	Y	Y		Y	Y	Y	Y

*Y: Yes, it was implemented

Fig. 2 Research design and procedure

list all words new to them on a worksheet,² look them up in an online dictionary, use their own words to explain the meaning of each word, and submit the worksheet together with the wind power passage at the end of the class.

The DMap group received a 15-min instructor-led presentation on how to argue with the DMap. It introduced argumentation concepts similar to those in the tutorial received by the comparison group, except they were presented in the context of the structure of a DMap. Like the comparison group, The DMap group was then given the wind power passage to read. After that, the participants were instructed to use the DMap to construct an argument map based on the information presented in the wind power passage to argue if wind power should be widely used. They were allowed to refer to the wind power passage while DMapping and include any ideas relevant to their arguments regardless of whether they appeared in the wind power passage. The participants were required to submit the DMap and hand in the wind power passage at the end of the class.

Session 2 was set up for the first delayed outcome assessment. Each participant was given 40 min to write an argument essay about wind power. They were provided with a copy of wind power passage to refer to while writing the essay. Although the DMap group had constructed an argument map on the same topic during the argumentation training session, they were not given access to the DMaps they created in order to maintain equal testing conditions for the two groups. After submitting the wind power essay, each group was given 15 min to complete the vocabulary test involving 15 definition questions.

Session 3 was set up for the second delayed outcome assessment. All participants were asked to write an essay arguing whether universities should increase use of computer technology for student learning. They were given 40 min to complete this essay, following which they were invited to respond to an online questionnaire specific to each group.

The 30-min QPT was conducted prior to session 1. All learning activities and assessments in the three sessions were completed in class with the same time allocated for the two groups. The posttest measures were paper-and-pencil tests graded for course credits.

To ensure treatment conditions were implemented, all learning activities and outcome tests were conducted in class and observed by the course instructor. The data from students who did not complete the learning activities or tests in class were excluded from this study.

3.6 Methods for data analysis

3.6.1 Quantitative data

The quantitative data were: (a) student demographic data consisting of age, gender, and English proficiency, (b) vocabulary test score, (c) essay performance data consisting of word count and the frequency of each argumentation component (pro argument, con argument, and rebuttal) in the wind power essay and the computer technology essay, and (d) responses to the Likert items in the argumentation and DMap questionnaires.

² This informed the selection of words to be included in the vocabulary test.

SPSS 27 was used to analyze the quantitative data. Independent samples *t*-tests were conducted to compare group differences. Prior to data analyses, several checks (e.g., outliers, normality, and homogeneity of variance) were carried out to ensure the assumptions for the statistical tests used in this study were met. The Mann–Whitney *U* test was used to compare group difference if the assumptions of normality and homogeneity of variance were not met. One participant in the comparison group who wrote 49 pro arguments ($z=6.89$) in the wind power essay was identified as an outlier and that pro argument score was adjusted to 24 using the recommendation of “one unit larger (or smaller) than the next most extreme score in the distribution” (Tabachnick et al., 2019, p. 67). Pearson correlation coefficients were used to investigate notable relationships among variables and inform selection of covariates. A one-way ANCOVA was conducted if the effect of a covariate was to be tested. Missing data were excluded test by test.

3.6.2 Qualitative data

The qualitative data were participant responses to the open-ended questions presented in the questionnaires. Microsoft Excel 2021 was used to facilitate analysis of qualitative data.

An inductive thematic analysis was conducted to identify the recurring patterns or key themes emerging from participants’ comments. Following Braun and Clarke (2006)’s six-step guidelines, after repeatedly reading through the verbal responses, codes were initially generated and collated into themes that were then refined and redefined iteratively.

4 Results

Table 2 presents the background information about participants in each group. An independent-samples *t* test showed the two groups did not differ significantly in age, $t(188)=1.41$, $p=0.16$. The DMap group had a greater proportion of female students than the comparison group. Given the correlation between gender and English proficiency tested prior to group-specific intervention was near zero (Comparison group: $r=-0.02$, $p=0.82$; DMap group: $r=0.09$, $p=0.42$), we assumed the unbalanced distribution of gender would have a negligible impact on the research findings.

There was a statistically detectable group difference in English proficiency measured by the QPT, $t(182)=2.07$, $p=0.04$, with the comparison group scoring higher than the DMap group. To inspect the potential impact of QPT achievement on the outcome variables, we examined the correlations among the QPT score, vocabulary

Table 2 Age, gender, and English proficiency of the two groups

	Age	Female	English Proficiency
Comparison Group	18.31 (0.51)	30.48%	34.49 (6.39)
DMap Group	18.21 (0.49)	48.24%	32.73 (4.89)

Means with standard deviations in parentheses

test performance, and all posttest essay component variables (i.e., pro argument, con argument, and rebuttal) for each group. Since the only variable significantly correlated with the QPT score was the vocabulary test score ($r=0.35$, $p<0.001$ for the comparison group and $r=0.22$, $p=0.04$ for the DMap group), the QPT variable was included as a covariate in the analyses of vocabulary test performance.

4.1 Analyses of vocabulary test performance

Excluding those who did not take the QPT or vocabulary test resulted in a sample of 84 participants in the DMap group ($M_{adjusted}=6.76$, $SE=0.35$) and 97 participants in the comparison group ($M_{adjusted}=5.91$, $SE=0.32$). A one-way ANCOVA was conducted to detect group difference in the adjusted mean for the vocabulary test after controlling for the QPT score, $F(1, 178)=3.15$, $p=0.08$, $\eta_p^2=0.02$. The proximity of the p -value to the conventional and arbitrary threshold for statistical significance suggests that a larger sample size might have resulted in a small but significant advantage ($d=0.27$) for the DMap group.

4.2 Analyses of argumentation components

Table 3 shows the number of participants in each group who submitted the wind power essay and the computer technology essay.

In addition to analyses of continuous data (e.g., the number of pro arguments included in each essay), binary coding was used to identify participants who included at least one pro argument, con argument, and rebuttal in each essay as a measure of awareness of using the key argumentation component(s).

4.2.1 Word count

Wind power essay For the wind power essay, an independent-samples t test revealed that the DMap group ($M=198.44$, $SD=37.95$) wrote statistically significantly longer essays than those in the comparison group ($M=164.51$, $SD=39.35$), $t(188)=6.00$, $p<0.001$, $d=0.88$.

Computer technology essay The DMap group ($M=180.67$, $SD=39.52$) and the comparison group ($M=171.69$, $SD=34.61$) wrote essays of similar length for the computer technology topic, $t(183)=1.65$, $p=0.10$, $d=0.24$.

Table 3 Number of essay submissions

	Wind Power Essay	Computer Technology Essay
Comparison Group	105	102
DMap Group	85	83

4.2.2 Pro argument

Wind power essay In the wind power essay, all participants gave pro arguments to defend the position taken. The DMap group wrote significantly more pro arguments ($M=13.99$, $SD=4.94$) than the comparison group ($M=12.56$, $SD=4.11$), $t(188)=2.17$, $p=0.03$, $d=0.31$.

Computer technology essay In the computer technology essay, all participants gave pro arguments to defend the position taken. The DMap group ($M=11.40$, $SD=4.13$) and the comparison group ($M=10.92$, $SD=3.63$) did not differ significantly in the number of pro arguments included in the computer technology essay, $t(183)=0.83$, $p=0.41$, $d=0.12$.

4.2.3 Con argument

Wind power essay The participants in the DMap group were more likely to present a counterargument in their argumentative essays. Specifically, 64.71% of participants in the DMap group and 54.29% of participants in the comparison group used counterarguments (i.e., included at least one con argument) in their wind power essays. A Mann–Whitney U test showed that the DMap group generated a greater number of con arguments (mean rank = 103.79) than the comparison group (mean rank = 88.79), $U=3758.00$, $z=-1.95$, $p=0.05$, $\eta^2=0.02$.

Computer technology essay For the computer technology essay, 61.45% of participants in the DMap group and 43.13% of participants in the comparison group were found to include the counterargument component. A Mann–Whitney U test revealed the DMap group included significantly more con arguments (mean rank = 107.09) than the comparison group (mean rank = 81.53) in the computer technology essay, $U=3063.50$, $z=-3.44$, $p<0.001$, $\eta^2=0.06$.

4.2.4 Rebuttal

Wind power essay For the wind power essay, 36.47% of participants in the DMap group and 33.33% of participants in the comparison group wrote at least one rebuttal. A Mann–Whitney U test indicated that the DMap group (mean rank = 97.21) and the comparison group (mean rank = 94.12) did not differ significantly in the number of rebuttals included in the wind power essay, $U=4317.50$, $z=-0.46$, $p=0.65$, $\eta^2=0.001$.

Computer technology essay For the computer technology essay, a greater proportion of participants in the DMap group (50.60%) presented a rebuttal than the comparison group (39.22%). However, a Mann–Whitney U test showed that there was no statistically significant difference between the DMap group (mean rank = 99.11) and the comparison group (mean rank = 88.98) in the number of rebuttals included in the computer technology essay, $U=3809.00$, $z=-1.42$, $p=0.16$, $\eta^2=0.01$.

Table 4 Proportion of essays involving three argumentation components

	Comparison Group	DMap Group
Wind Power Essay	31.40%	36.47%
Computer Technology Essay	38.24%	50.60%

Taken together, a higher percentage of participants in the DMap group showed awareness of using all three components of argumentation (i.e., including at least one proposition in each category of pro argument, con argument, and rebuttal), especially in their computer technology essays that measured learning transfer (see Table 4).

4.3 Correlations

Correlations among posttest essay component measures were conducted for each group. An interesting finding was the comparison group did not show a statistically detectable relationship between the number of con arguments presented in the wind power essay and the number of con arguments included in the computer technology essay ($r=0.14$, $p=0.16$). However, their relation was detectable in the DMap group ($r=0.22$, $p=0.04$). The same pattern applied to the relationship between the number of rebuttals provided in each essay ($r=0.11$, $p=0.27$ for the comparison group and $r=0.28$, $p=0.01$ for the DMap group), indicating a more robust transfer effect of DMapping on argumentation.

4.4 DMap questionnaire responses

The participants ($n=86$) in the DMap group responded to a questionnaire involving 28 Likert items. Each statement had five options indicating different levels of agreement ranging from *strongly disagree* to *strongly agree*, assigned 1 to 5 points, respectively. Table 5 shows the mean score associated with each factor identified in Deng et al. (2021), suggesting positive experience and attitudes toward studying with the DMap tool.

Table 5 DMap questionnaire factors and mean scores

Factor	Mean Score
Learning and Motivation <i>Example: Using the DMap motivates me to learn argumentation</i>	3.70
Self-Efficacy <i>Example: I feel confident using the DMap</i>	3.23
Instructional Mode <i>Example: Instructors should use DMapping to teach argumentation</i>	3.52
Usability <i>Example: The DMap interface worked the way I wanted it to work</i>	3.61

4.5 Argumentation questionnaire responses

There were 98 participants in the comparison group and 83 in the DMap group who responded to the argumentation questionnaire involving three 5-point Likert items about the structure of argumentation and two open-ended questions investigating why students included counterarguments in each essay or not.

We found 86.73% of participants in the comparison group and 75.90% in the DMap group responded *Agree* or *Strongly Agree* to the statement that discussing counterarguments strengthens one's essay. Also, 96.94% of participants in the comparison group and 79.52% in the DMap group responded *Agree* or *Strongly Agree* to the statement that one should consider both sides of an argument when developing a conclusion. The other item was negatively worded, with 65.31% of participants in the comparison group and 73.49% in the DMap group responding *Disagree* or *Strongly Disagree* to the statement that an argument essay should only present one side of an argument.

Table 6 presents the mean score given by participants in each group, indicating the intervention activities received by both groups were effective in helping participants acquire basic ideas about the structure of argumentation.

4.6 Explanations about counterargumentation

In response to open-ended questions about their use of counterargumentation, the groups provided similar explanations. Those who reported using counterarguments in their essays mostly referred to the tutorials that highlighted the importance of presenting arguments on both sides to more convincingly defend the position taken. In the same vein, a few students in the DMap group specifically mentioned the structure of the DMap used in argumentation training helped as it directed their attention to and prompted them to develop counterarguments. The other recurring explanation was that information in the reading passage cued them to include counterarguments in the wind power essays, for example:

I found there were several sentences in the wind power text discussing the disadvantages of using wind power and wind farms. I feel I should say something about it as well in my essay because the instructor asked us to argue based on the wind power passage we read.

There were five frequently mentioned themes explaining why counterarguments were not presented, including (1) lack of awareness of making counterarguments, (2) lack of skill in counterargumentation, (3) time limit, (4) lack of prior knowledge of subject, and (5) essay length. All factors, except prior knowledge of the topic to argue, seem to be tied to a flawed argument schema that directs one's argumentation moves and planning. See Table 7 for further explanations and sample comments.

Table 6 Argumentation questionnaire: mean scores

Likert Item	Comparison Group	DMap Group
Discussing counterarguments (i.e., the arguments of an opponent) strengthens my essay	4.11	3.78
*An argument essay should only present one side of the argument	2.33	2.14
You should consider both sides of an argument when you develop a conclusion	4.04	3.93

*Negatively worded

Table 7 Frequently mentioned reasons for not using counterarguments

Reason	Explanation	Sample Comments
Lack of Awareness of Making Counterarguments	Inadequate awareness of the need for stating counterarguments or discussing opposing perspectives in an essay to defend the position one takes	<i>I forgot to include [cons] because I seldom consider that part when I write essays. Now I remember the instructor did say we should include [cons] in our essays to better support our viewpoint</i> <i>No, we were asked to select one option most consistent with what we believe, so those arguments that I don't really believe were ignored</i>
Lack of Skill in Counterargumentation	Not skilled in effectively integrating counterarguments into an essay for enhanced outcomes	<i>I don't know how to add a counterargument. It feels like adding cons makes my essay less coherent</i> <i>I don't know how to connect the cons to my claim and how to make use of cons to strengthen my arguments</i>
Time Limit	Insufficient time left for coming up with counterarguments, which might indicate ineffective planning for argumentative writing	<i>There was no time left for cons</i> <i>I didn't want to spend too much time [thinking of cons] because I'm afraid I wouldn't be able to finish the essay</i>
Lack of Prior Knowledge of Subject	Unable to develop pertinent counterarguments due to limited knowledge about the topic to argue	<i>I didn't figure out any cons [of using computer technology in education]</i> <i>I don't know if the cons I wanted to write are true. I don't want to take the risk of including incorrect information in my essay</i>
Essay Length	No perceived need for including counterarguments as the essay has already met the length requirement, which might imply a misconception about the purpose of including counterarguments in an argument essay	<i>The essay has already been long. There is no space for counterarguments</i> <i>We just need to write an essay over 150 words. I've got enough words so there is no need to write more</i>

Since time, topic, and word limit for the essay tests were the same for both groups, the greater counterargumentation in the DMap group can be mainly attributed to the DMapping treatment developing an enhanced awareness of and increased skill in integrating counterarguments into essays.

5 Discussion

5.1 Effects of DMapping as a pre-writing activity for argumentative writing

The findings of this study indicated that DMapping, designed as a constructive approach to learning argumentation, could be considered an effective addition to an instructor-led presentation of argument structure. Engagement in DMapping as a pre-writing intervention helped students write significantly lengthier essays with more pro arguments and counterarguments than participants who, in addition to the instructor-led presentation of argument structure, completed a non-argumentative pre-writing activity focused on understanding the vocabulary used in the given text passage. This result is in line with that of previous research that mere presentation of information is insufficient to equip students with advanced argumentation skills (Niu, 2016; Nussbaum, 2008). Unlike the teacher-led presentations of argument concepts received by both groups, scaffolded practice of DMapping fosters a constructive mode of cognitive engagement (Chi, 2009; Chi & Wylie, 2014) and can more effectively develop functional argument schemas that guide students' construction of balanced, well-founded arguments (Liu, 2020; Nesbit et al., 2019).

The two groups did not differ in the number of rebuttals presented in their essays probably because this component was not specifically emphasized with examples in the training tutorial received by either group. Even though one DMap feature enables students to link directly related pro and con arguments, the interface provides few visual signals that would cue use of this feature to refute an opposing argument. This observation points to areas of improvement for the DMap tool and underscores the importance of providing explicit instructions and specialized training in argumentation tutorials to foster acquisition of the key argumentation skill of rebuttal (Liu, 2020; Berland & Reiser, 2009; Jiménez-Aleixandre & Erduran, 2007; Niu, 2016). Taking another perspective, the interactive mode of the ICAP framework suggests that a more direct route to developing counterargumentation and refutation could be opened by adding a collaborative argumentation feature to the DMap tool.

5.2 Effects of DMapping on transfer of argumentation skills

The computer technology essay test that took place two weeks after the learning session measured transfer of argumentation skills. The participants were not exposed to any pre-writing interventions or provided with any reference readings that might potentially interfere with how they argue. It was found that a larger proportion of participants in the DMap group discussed and refuted opposing perspectives in their computer technology essays. The Dmappers who used con

arguments and rebuttals in the wind power essay were more likely to include those argumentation components in the computer technology essay, but the correlations were not measurable for the comparison group, indicating the effect of DMapping on argumentation was more robust or transferable than the argumentation training received by the comparison group.

Looking at the overall frequency of each argumentation component generated in the transfer essay, we found that the two groups did not differ in the number of words, pro arguments and rebuttals, but the DMap group produced significantly more counterarguments than the comparison group. We assumed that all participants had more experience or prior knowledge about the use of educational technology than about wind power. If this is the case, given the key role of subject matter knowledge in argumentation, the participants in both groups might be able to readily self-generate abundant related arguments without exposure to additional prompts (Cross et al., 2008; Wang & Buck, 2015). Consequently, group differences in word count and pro argument were not as salient as those reflected in the wind power essay that likely requires more effortful processing of incoming information for argument construction. In other words, DMapping might not always be helpful. The beneficial effect of studying with the DMap as a cognitive tool may become more pronounced in situations where students have less prior knowledge and consequently experience higher cognitive load induced by processing and operating on information from multiple sources.

The efficacy of the DMap can also be explained by the three essential attributes of cognitive tools proposed by Pakdaman-Savoji et al. (2019). The DMap visually represents the major constituents of argumentation and the relationships among them. When students are searching for and generating ideas, the DMap serves as an externally distributed memory they can refer to for a current overview of which components and connections have been completed and which are lacking. By interacting with the tool on multiple occasions, students can gradually refine their argument schema until they no longer need it to remind them how to construct dialectical arguments.

In a similar study by Nussbaum (2008), the effect of argument mapping was not transferred to a new argument essay constructed without the aid of a vee diagram. The present research, however, detected a more robust transfer effect of DMapping on argumentation. A comparison of the AVTs used (e.g., ease of revision and visual prompts for different argument components and processes) might shed light on the differing outcomes.

5.3 Effects of DMapping on vocabulary acquisition

Unlike those in the comparison group, the participants in the DMap group were not specifically instructed to study the vocabulary in the text passage but they did as well as the comparison group in the vocabulary test that took place one week after the learning session.

Given that the comparison group engaged in a similar vocabulary exercise at the intervention stage, the result to some extent indicates there is a potential tension

between the transfer-appropriate processing assumption (Morris et al., 1977) and the levels of processing theory (Craik & Lockhart, 1972) in this context. This research lends support to the latter and accentuates an interpretation that argument mapping induces deep processing of information to be learned and fosters meaningful learning (Chi et al., 2018; Liu, 2020; Nesbit et al., 2019).

5.4 Student beliefs and attitudes toward DMapping

Based on the responses to the 28-item DMap questionnaire, it was found that the participants in the DMap group showed somewhat positive attitudes toward DMapping expressed as belief that DMapping motivates learning, preferences for DMapping as an instructional mode, and confidence (self-efficacy) in their DMapping ability. They also expressed the belief that the DMap interface is easy to use. A majority of participants agreed to statements suggesting instructors employ the DMap to teach and assess argumentation.

Among the four factors, the one that received the lowest mean score was self-efficacy. Since the participants rated the DMap to be easy to use, their low self-efficacy was likely due to limited exposure to the DMap tool for practice. In this study, the participants only used the DMap once before the posttests. Repeated task-related practice leads to enactive mastery experience that underlies the development of self-efficacy (Bandura, 1977). Another possibility was that the DMaps created by the participants were not graded, and the lack of feedback may have attenuated their self-efficacy in arguing with the DMap.

5.5 Reasons for inclusion of counterarguments

The argumentation questionnaire responses indicated that tutorials played a crucial role in acquisition of argumentation knowledge. About a third of participants claimed that their use of argumentation components was directed by what they learned from the argumentation tutorials. One interesting finding gleaned from the questionnaire data was both groups recognized the indispensable role of counterarguments, but they performed differently in the essay tests. Specifically, fewer participants in the comparison group took opposing perspectives into account and included a significantly smaller number of counterarguments in their essays. Griffin (1998) contended active cognitive engagement (e.g., clear thinking, careful planning, and deep reflection) bridges the gap between knowing and doing, leading to successful writing performance. De Corte (2003) maintained that powerful learning environments that yield successful transfer allow ample opportunities for practice and productive use of acquired knowledge, skills, and motivations. This may have explained why the DMap group outperformed the comparison group in the use of counterargument component and were better able to bridge the knowing-doing gap.

According to the open-ended responses, the most frequently mentioned explanations for not including counterarguments in their essays were: (1) lack of awareness

of the necessity for stating counterarguments, (2) lack of skill in counterargumentation, (3) time limit, (4) lack of prior knowledge about the topic to argue, and (5) essay length. Except subject domain knowledge, all other factors seem to be associated with a deficient argument schema underlying one's argumentation moves and planning. For the essay tests in this study, the latter three concerns (i.e., time, essay topic, and length) were considered equivalent across groups. The difference between treatment groups in the number of counterarguments might be best explained by the first two reasons relating to awareness and skill in developing counterarguments. Compared to instructor-led tutorials, coupling an argumentation tutorial with DMapping seemed to be more effective in strengthening students' awareness of attending to counterarguments and improving their skills in integrating counterarguments into argumentative essays for enhanced outcomes.

5.6 Implications

Previous studies, which were conducted in laboratory settings with fully randomized control, reported the effectiveness of DMapping in promoting learning and argumentation (Liu, 2020; Nesbit et al., 2019; Niu, 2016). The present study was conducted in a classroom setting to investigate if the efficacy of DMapping is generalizable to authentic learning environments outside the experimental setting. Taken together, these laboratory and classroom studies support the pedagogical value of DMapping.

Prior research has demonstrated the DMap could be used to facilitate learning in different contexts. Niu (2016) asked participants to construct DMaps while studying a scientific text, leading to better recall of main ideas. Liu (2020) carried out a lab experiment with students from a university in Canada. After the reading session, they were instructed to recall all information presented in the text and create DMaps based on what they could remember. The findings revealed the efficacy of DMapping as retrieval practice in fostering learning transfer (i.e., applying learned concepts or skills to solve novel problems). Building on previous research, the present study was conducted with Chinese EFL learners in authentic postsecondary classrooms. It was found that DMapping could be implemented as an effective pre-writing activity to supplement the tutorial on argumentation. The participants conveyed positive experience and attitudes toward arguing with the DMap. In addition to argumentative writing, DMapping contributed to incidental vocabulary learning. These findings justify more extensive use of DMapping in postsecondary education and further investigation of its potential for promoting learning outcomes across the curriculum.

Pakdaman-Savoji et al. (2019) held that repeated engagement with a cognitive tool obtains more robust effects than a single usage. Theoretically, iterative interaction with the DMap promotes development of a corresponding argument schema (i.e., internalization) that when fully formed enables students to argue effectively without external prompts (Nesbit et al., 2019). In light of the technology acceptance model, an individual's intention to use a new technology is influenced by two factors: perceived ease of use and perceived usefulness (Davis, 1989). One's

self-efficacy influences perceived usefulness of a new educational tool or environment (Liaw & Huang, 2013). In order to encourage frequent use of the DMap and maximize its efficacy in advancing learning and argumentation, provision of sufficient training and positive learning experiences reinforced through instructional support and feedback would seem to be crucial.

5.7 Limitations and future research

This study was conducted in an authentic classroom setting, so it bears the threats to internal validity typically related to quasi-experimental designs such as pre-existing group differences and unequal class sizes that were beyond the researchers' control. After receiving the instructor-led presentation of argument structure, the comparison group did not engage in any argumentative pre-writing activities but focused on understanding the vocabulary in the reading passage provided to support argument construction. Therefore, the research design only examined the effect of the DMap treatment relative to a non-argumentative, content-relevant learning activity.

Scheduling and curricular constraints meant there was no opportunity to assess the participants' prior knowledge and skill in argumentation. Even though the instructor teaching the four course sections from which the participants were drawn commented that most students performed poorly in argumentative tasks, the research would benefit from direct measurement of this variable to investigate if it moderates the effect of DMapping on enhancing argumentative writing. Also, it would be worthwhile investigating if repeated learning activities using the DMap would produce different or more robust cognitive effects.

Compared to other commonly used AVTs like *Rationale*TM and argumentation vee diagram, the DMap provides more scaffolding for advanced argumentation. For example, it prompts learners to elaborate on how evidence supports or opposes the reason it associates with (i.e., warrant), (re)sequence the arguments for rhetorical purposes, and evaluate the strength of each argument as the basis for making a well-reasoned conclusion (Nesbit et al., 2019; Niu, 2016). However, the DMap tool may also have disadvantages. As mentioned above, it appears to be less effective in eliciting rebuttals and lacks features that would support collaborative engagement. The bilateral structure of the DMap is easier to read, but it may at the same time restrict the way more sophisticated arguers present their reasoning. Given their strengths and limitations, we suggest future research directly compare the efficacies of AVTs in promoting a defined learning outcome to illuminate the critical features or affordances that affect learning and argumentation in different contexts.

In this study, the DMaps created during the intervention phase were not analyzed. In future investigations, it would be of interest to compare the similarities between the DMap and the argument essay it corresponds to, which to some extent reveals how transfer of argumentation skills occurs.

von Aufschnaiter et al. (2008) contended that the quality of argumentation is affected by students' experience and familiarity with the topic to argue. It is therefore important to examine if argumentation topic interacts with the effects of DMapping.

In this study, the participants engaged in an argumentation task on a topic (i.e., wind power) less familiar to many of them at the intervention phase. If students are asked to argue with the DMap on a topic they know well (e.g., impact of COVID-19) as the research treatment, they may not need to put much effort into processing the reading material and could spare more cognitive capacity for argument analysis and construction scaffolded by the DMap, which would likely induce more robust learning effects.

To strengthen the foundation for designing instructional experiences conducive to development of argumentation skills and better understand the factors that impact how students learn to argue, it is necessary to replicate and extend this research with larger and more varied samples.

Acknowledgements We are grateful to the following persons for their invaluable assistance: Arati Santhanakrishnan, Lu Zhou, Yaling Zhu, Jikun Qian, Yanling Zhong, Zhiling Wu, Xiaohong Chen, Yating Li, Wenjue Yu, and Han Liu.

Authors' contributions

- The first author (Qing Liu) conducted the research design, created the research materials and instruments, assisted in data collection, analyzed the research data, and drafted the manuscript.
- The second author (Zhiying Zhong) provided suggestions on research design, conducted data collection, and assisted in data analysis.
- The third author (John C. Nesbit) provided suggestions on research design and data analysis and edited the manuscript.

Funding This research was funded by the Educational Bureau of Guangdong Province, PRC (2019GXJK162).

Data availability The datasets and materials used in the present study are available from the corresponding author upon request.

Declarations

Competing interests None.

References

- Abdelaal, M., Schiele, N. D., Angerbauer, K., Kurzhals, K., Sedlmair, M., & Weiskopf, D. (2022). Comparative evaluation of bipartite, node-link, and matrix-based network representations. *IEEE Transactions on Visualization and Computer Graphics*, 29(1), 896–906.
- Adesope, O. O., & Nesbit, J. C. (2013). Animated and static concept maps enhance learning from spoken narration. *Learning and Instruction*, 27, 1–10.
- Al-Shaer, I. M. (2014). Employing concept mapping as a pre-writing strategy to help EFL learners better generate argumentative compositions. *International Journal for the Scholarship of Teaching and Learning*, 8(2), 1–29.
- Anderson, R. C., Nguyen-Jahiel, K., McNurlen, B., Archodidou, A., Kim, S. Y., Reznitskaya, A., Tillmanns, M., & Gilbert, L. (2001). The snowball phenomenon: Spread of ways of talking and ways of thinking across groups of children. *Cognition and Instruction*, 19(1), 1–46.
- Andrews, R. (2015). Critical thinking and/or argumentation in higher education. In M. Davies & R. Barnett (Eds.), *The Palgrave handbook of critical thinking in higher education* (pp. 49–62). Palgrave Macmillan.
- Bacha, N. N. (2010). Teaching the academic argument in a university EFL environment. *Journal of English for Academic Purposes*, 9(3), 229–241.

- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84(2), 191–215.
- Beardsley, M. C. (1950). *Practical logic*. Prentice-Hall.
- Berland, L. K., & Reiser, B. J. (2009). Making sense of argumentation and explanation. *Science Education*, 93(1), 26–55.
- Binks, A., Toniolo, A., & Nacenta, M. A. (2022). Representational transformations: Using maps to write essays. *International Journal of Human-Computer Studies*, 165(102851), 1–20.
- Bovair, S., & Kieras, D. E. (1985). A guide to propositional analysis for research on technical prose. In B. K. Britton & J. B. Black (Eds.), *Understanding expository text* (pp. 315–362). Erlbaum.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3, 77–101.
- Burch, M., Huang, W., Wakefield, M., Purchase, H. C., Weiskopf, D., & Hua, J. (2020). The state of the art in empirical user evaluation of graph visualizations. *IEEE Access*, 9, 4173–4198.
- Chi, M. T. H. (2009). Active-constructive-interactive: A conceptual framework for differentiating learning activities. *Topics in Cognitive Science*, 1(1), 73–105.
- Chi, M. T., & Wylie, R. (2014). The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational Psychologist*, 49(4), 219–243.
- Chi, M. T., Adams, J., Bogusch, E. B., Bruchok, C., Kang, S., Lancaster, M., ... & Yaghmourian, D. L. (2018). Translating the ICAP theory of cognitive engagement into practice. *Cognitive Science*, 42(6), 1777–1832.
- Chiang, K. H., Fan, C. Y., Liu, H. H., & Chen, G. D. (2016). Effects of a computer-assisted argument map learning strategy on sixth-grade students' argumentative essay reading comprehension. *Multimedia Tools and Applications*, 75(16), 9973–9990.
- Craik, F. I., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 11(6), 671–684.
- Cross, D., Taasoobshirazi, G., Hendricks, S., & Hickey, D. T. (2008). Argumentation: A strategy for improving achievement and revealing scientific identities. *International Journal of Science Education*, 30(6), 837–861.
- Davies, M., Barnett, A., & van Gelder, T. (2021). Using computer-assisted argument mapping to teach reasoning to students. In J. Anthony Blair (Ed.), *Studies in critical thinking* (pp. 115–152), Windsor Studies in Argumentation (2021), Windsor, ON.
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *Management Information Systems Quarterly*, 13(3), 319–340.
- De Corte, E. (2003). Transfer as the productive use of acquired knowledge, skills, and motivations. *Current Directions in Psychological Science*, 12(4), 142–146.
- Deng, Y., & Wang, H. (2017). Research on evaluation of Chinese students' competence in written scientific argumentation in the context of chemistry. *Chemistry Education Research and Practice*, 18(1), 127–150.
- Deng, Q. Y., Nesbit, J. C., Liu, Q., Sharp, J., Cukierman, D., Shahabi, B., Teng, K., Pakdaman Savoji, A., Ilten-Gee, R., & Flores, O. (2021). Assessing Student Attitudes Toward an Argument Visualization Tool [Paper presentation]. The 4th International Conference on Teaching, Learning and Education (ICTLE), Zurich, Switzerland.
- Dole, J. A., & Sinatra, G. M. (1998). Reconceptualizing change in the cognitive construction of knowledge. *Educational Psychologist*, 33(2/3), 109–128.
- Dwyer, C. P., Hogan, M. J., & Stewart, I. (2012). An evaluation of argument mapping as a method of enhancing critical thinking performance in e-learning environments. *Metacognition and Learning*, 7(3), 219–244.
- Dwyer, C. P., Hogan, M. J., & Stewart, I. (2013). An examination of the effects of argument mapping on students' memory and comprehension performance. *Thinking Skills and Creativity*, 8, 11–24.
- El Majidi, A., Janssen, D., & de Graaff, R. (2021). The effects of in-class debates on argumentation skills in second language education. *System*, 101(102576), 1–15.
- Fleckenstein, J., Meyer, J., Jansen, T., Keller, S., & Köller, O. (2020). Is a long essay always a good essay? The effect of text length on writing assessment. *Frontiers in Psychology*, 11, 562462. <https://doi.org/10.3389/fpsyg.2020.562462>
- Garcia Romano, L., Occelli, M., & Adúriz-Bravo, A. (2021). School scientific argumentation enriched by digital technologies: Results with pre-and in-service science teachers. *Eurasia Journal of Mathematics, Science and Technology Education*, 17(7), 1–16.

- Ghanbari, N., & Salari, M. (2022). Problematising argumentative writing in an Iranian EFL undergraduate context. *Frontiers in Psychology, 13*(862400), 1–11.
- Glassner, A., & Schwarz, B. B. (2007). What stands and develops between creative and critical thinking? Argumentation? *Thinking Skills and Creativity, 2*(1), 10–18.
- Griffin, C. W. (1998). Improving students' writing strategies: Knowing versus doing. *College Teaching, 46*(2), 48–52.
- Hege, I., Kononowicz, A. A., & Adler, M. (2017). A clinical reasoning tool for virtual patients: Design-based research study. *JMIR Medical Education, 3*(2), 1–11.
- Heitmann, P., Hecht, M., Schwanewedel, J., & Schipolowski, S. (2014). Students' argumentative writing skills in science and first-language education: Commonalities and differences. *International Journal of Science Education, 36*(18), 3148–3170.
- Hogan, K., & Fisherkeller, J. (2000). Dialogue as data: Assessing students' scientific reasoning with interactive protocols. In J. J. Mintzes, J. H. Wandersee, & J. D. Novak (Eds.), *Assessing science understanding: A human constructivist view* (pp. 95–127). Academic.
- Huang, W. (2007). Using eye tracking to investigate graph layout effects. In *2007 6th International Asia-Pacific Symposium on Visualization* (pp. 97–100). IEEE.
- Iordanou, K., & Rapanta, C. (2021). “Argue with me”: A method for developing argument skills. *Frontiers in Psychology, 12*(631203), 1–14.
- 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.
- Jiménez-Aleixandre, M. P., & Puig, B. (2012). Argumentation, evidence evaluation and critical thinking. In B. Fraser, K. Tobin, & C. McRobbie (Eds.), *Second international handbook of science education* (pp. 1001–1015). Springer.
- Jo, C. W. (2021). Exploring general versus academic English proficiency as predictors of adolescent EFL essay writing. *Written Communication, 38*(2), 208–246.
- Johnson, B., & Christensen, L. (2008). *Educational research: Quantitative, qualitative, and mixed approaches*. SAGE. ISBN 9781412954563.
- Jonassen, D. H., & Kim, B. (2010). Arguing to learn and learning to argue: Design justifications and guidelines. *Educational Technology Research and Development, 58*(4), 439–457.
- Jonassen, D., & Reeves, T. (1996). Learning with technology: Using computers as cognitive tools. In D. Jonassen (Ed.), *Handbook of research on educational communication and technology* (pp. 693–719). Macmillan.
- Jumariati, J., Febriyanti, E., & Rizki, M. (2021). Argumentation skills: An analysis on EFL students' essays based on Toulmin's Model of Argument. *Proceedings of the 2nd International Conference on Education, Language, Literature, and Arts*, 86–95.
- Khartabil, D., Collins, C., Wells, S., Bach, B., & Kennedy, J. (2021). Design and evaluation of visualization techniques to facilitate argument exploration. *Computer Graphics Forum, 40*(6), 447–465.
- Kuhn, D., & Crowell, A. (2011). Dialogic argumentation as a vehicle for developing young adolescents' thinking. *Psychological Science, 22*(4), 545–552.
- Kuhn, D., Shaw, V., & Felton, M. (1997). Effects of dyadic interaction on argumentative reasoning. *Cognition and Instruction, 15*(3), 287–315.
- Lajoie, S. (1993). Computer environments as cognitive tools for enhancing learning. In S. Lajoie & S. Derry (Eds.), *Computers as cognitive tools* (pp. 261–288). Lawrence Erlbaum Associates.
- Liao, M., & Liao, Y. (2022). Improving Chinese college students' argumentative writing: A presentation-assimilation-discussion-exercise approach. *Frontiers in Psychology, 13*, 874531. <https://doi.org/10.3389/fpsyg.2022.874531>
- Liaw, S. S., & Huang, H. M. (2013). Perceived satisfaction, perceived usefulness and interactive learning environments as predictors to self-regulation in e-learning environments. *Computers & Education, 60*(1), 14–24.
- Liu, Q. (2020). *Retrieval-based argument mapping promotes learning transfer* (Unpublished doctoral dissertation), Simon Fraser University, Burnaby, BC, Canada.
- Luo, X., Wei, B., Shi, M., & Xiao, X. (2020). Exploring the impact of the reasoning flow scaffold (RFS) on students' scientific argumentation: Based on the structure of observed learning outcomes (SOLO) taxonomy. *Chemistry Education Research and Practice, 21*(4), 1083–1094.

- Mejía, K. P. G., & Neve, L. J. A. (2018). The problem of argumentation teaching in Mexican schools (El problema de la enseñanza de la argumentación en la escuela Mexicana). *Pensamiento Educativo, Revista De Investigación Educativa Latinoamericana*, 55(2), 1–18.
- Mitchell, S., & Riddle, M. (2000). Improving the quality of argument in higher education: Final report. *School of Lifelong Learning and Education, Middlesex University*.
- Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior*, 16(5), 519–533.
- National Research Council (U.S.). (1996). *National science education standards: Observe, interact, change, learn*. National Academy Press.
- Nesbit, J., Niu, H., & Liu, Q. (2019). Cognitive tools for scaffolding argumentation: Maximizing student engagement, motivation, and learning. In O. Adesope & A. Rud (Eds.), *Contemporary technologies in education* (pp. 97–117). Palgrave Macmillan.
- Ngajie, B. N., Li, Y., Tiruneh, D. T., & Cheng, M. (2020). Investigating the effects of a systematic and model-based design of computer-supported argument visualization on critical thinking. *Thinking Skills and Creativity*, 38(100742), 1–15.
- Niu, H. (2016). *Pedagogical efficacy of argument visualization tools* (Unpublished doctoral dissertation), Simon Fraser University, Burnaby, BC, Canada.
- Noroozi, O. (2023). The role of students' epistemic beliefs for their argumentation performance in higher education. *Innovations in Education and Teaching International*, 60(4), 501–512.
- Nussbaum, E. M. (2008). Using argumentation vee diagrams (AVDs) for promoting argument-counterargument integration in reflective writing. *Journal of Educational Psychology*, 100(3), 549–565.
- Nussbaum, E. M. (2011). Argumentation, dialogue theory, and probability modeling: Alternative frameworks for argumentation research in education. *Educational Psychologist*, 46(2), 84–106.
- Nussbaum, E. M., & Schraw, G. (2007). Promoting argument-counterargument integration in students' writing. *Journal of Experimental Education*, 76(1), 59–92.
- Ogan-Bekiroglu, F., & Eskin, H. (2012). Examination of the relationship between engagement in scientific argumentation and conceptual knowledge. *International Journal of Science and Mathematics Education*, 10(6), 1415–1443.
- Osman, W. H., & Januin, J. (2021). Analysing ESL persuasive essay writing using Toulmin's model of argument. *Psychology and Education*, 58(1), 1810–1821.
- Oxford University Press. (2001). *Quick placement test*. Oxford University Press.
- Pakdaman-Savoji, A., Nesbit, J. C., & Gajdamaschko, N. (2019). The conceptualisation of cognitive tools in learning and technology: A review. *Australasian Journal of Educational Technology*, 35(2), 1–24.
- Pea, R. D. (1985). Beyond amplification: Using the computer to reorganize mental functioning. *Educational Psychologist*, 20(4), 167–182.
- Pei, Z., Zheng, C., Zhang, M., & Liu, F. (2017). Critical thinking and argumentative writing: Inspecting the association among EFL learners in China. *English Language Teaching*, 10(10), 31–42.
- Reznitskaya, A., & Anderson, R. (2002). The argument schema and learning to reason. In C. Block & M. Pressley (Eds.), *Comprehension instruction: Research-based best practices* (pp. 319–334). The Guilford Press.
- Reznitskaya, A., Anderson, R., McNurlen, B., Nguyen-Jahiel, K., Archodidou, A., & Kim, S. (2001). Influence of oral discussion on written argument. *Discourse Processes*, 32(2–3), 155–175.
- Rumelhart, D. E., & Norman, D. A. (1976). *Accretion, tuning and restructuring: three modes of learning*. University of California.
- Salomon, G., Perkins, D. N., & Globerson, T. (1991). Partners in cognition: Extending human intelligence with intelligent technologies. *Educational Researcher*, 20(3), 2–9.
- Sawyer, R. (2014). Introduction: The New Science of Learning. In R. Sawyer (Ed.), *The Cambridge Handbook of the Learning Sciences* (Cambridge Handbooks in Psychology, pp. 1–18). Cambridge University Press.
- Scriven, M. (1976). *Reasoning*. McGraw-Hill.
- Stapleton, P. (2011). A survey of attitudes towards critical thinking among Hong Kong secondary school teachers: Implications for policy change. *Thinking Skills and Creativity*, 6(1), 14–23.
- Tabachnick, B., Fidell, L., & Ullman, J. (2019). *Using multivariate statistics* (7th ed.). Pearson.
- Tan, C. (2020). Conceptions and practices of critical thinking in Chinese schools: An example from Shanghai. *Educational Studies*, 56(4), 331–346.
- Thomas, S. N. (1973). *Practical reasoning in natural language*. Prentice-Hall.
- Toulmin, S. (1958). *The uses of argument*. Cambridge University Press.

- van den Braak, S. W., Oostendorp, H. V., Prakken, H., & Vreeswijk, G. A. (2008). A critical review of argument visualization tools: Do users become better reasoners?. In Workshop notes of the ECAI-06 workshop on computational models of natural argument (pp. 67–75).
- van Gelder, T. (2007). The rationale for Rationale™. *Law, Probability and Risk*, 6(1–4), 23–42.
- van Gelder, T. (2015). Using argument mapping to improve critical thinking skills. In M. Davies & R. Barnett (Eds.), *The Palgrave handbook of critical thinking in higher education* (pp. 183–192). Palgrave Macmillan.
- van Eemeren, F., Grootendorst, R., & Henkemans, F. (1996). *Fundamentals of argumentation theory: A handbook of historical backgrounds and contemporary developments*. L. Erlbaum.
- van Eemeren, F. H., Grootendorst, R., & Meuffels, B. (2015). The skill of identifying argumentation. In F. H. van Eemeren (Ed.), *Reasonableness and effectiveness in argumentative discourse* (pp. 733–741). Springer.
- 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: The Official Journal of the National Association for Research in Science Teaching*, 45(1), 101–131.
- Walton, D. (2013). *Methods of argumentation*. Cambridge University Press.
- Wang, J., & Buck, G. (2015). The relationship between Chinese students' subject matter knowledge and argumentation pedagogy. *International Journal of Science Education*, 37(2), 340–366.
- Westrick, P. A. (2005). Score reliability and placement testing. *Japan Association for Language Teaching*, 27(1), 71–92.
- Whately, R. (1834/1826). *Elements of Logic: Comprising the Substance of the Article in the Encyclopaedia Metropolitana*. B. Fellowes.
- Wigmore, J. (1913). The problem of proof. *Illinois Law Review*, 8(2), 77–103.
- Wigmore, J. (1937). *Science of judicial proof, as given by logic, psychology, and general experience and illustrated in judicial trials*. Little, Brown and Co.
- Wingate, U. (2012). 'Argument!' helping students understand what essay writing is about. *Journal of English for Academic Purposes*, 11(2), 145–154.
- Wolfe, C. R., & Britt, M. A. (2008). The locus of the myside bias in written argumentation. *Thinking & Reasoning*, 14(1), 1–27.
- Wolfe, C. R., Britt, M. A., & Butler, J. A. (2009). Argumentation schema and the myside bias in written argumentation. *Written Communication*, 26(2), 183–209.
- Xie, Q., & So, W. (2012). Understanding and practice of argumentation: A pilot study with Mainland Chinese pre-service teachers in secondary science classrooms. *Asia-Pacific Forum on Science Learning and Teaching*, 13(2), 1–20.
- Yamamoto, T., & Kamiyama, S. (2018). Teacher beliefs about argumentation in Japanese in-service teachers. In O. Finlayson, E. McLoughlin, S. Erduran, & P. Childs (Eds.), *Electronic Proceedings of the ESERA 2017 Conference: Research, Practice and Collaboration in Science Education* (pp. 952–959). Dublin City University.
- Yamamoto, T., Yamaguchi, E., Muratsu, K., Nakashin, S., & Inagaki, S. (2014). Evaluation of Japanese pre-service teachers' arguments on global warming. In C. P. Constantinou, N. Papadouris, & A. Hadjigeorgiou (Eds.), *E-book proceedings of the ESERA 2013 conference* (pp. 2284–2289). ESERA.
- Yoghourdjian, V., Yang, Y., Dwyer, T., Lawrence, L., Wybrow, M., & Marriott, K. (2021). Scalability of network visualisation from a cognitive load perspective. *IEEE Transactions on Visualization and Computer Graphics*, 27(2), 1677–1687.
- Zhao, G., Zhao, R., Li, X., Duan, Y., & Long, T. (2023). Are preservice science teachers (PSTs) prepared for teaching argumentation? Evidence from a university teacher preparation program in China. *Research in Science & Technological Education*, 41(1), 170–189.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.