



Developing argumentation skills for democratic citizenship: addressing controversial societal issues with personal stakes

Kalypso Iordanou^{a,*}, Constantina Fotiou^{a,b}, Michalinos Zembylas^b

^a University of Central Lancashire, Cyprus

^b Open University Cyprus, Cyprus

ARTICLE INFO

Keywords:

Argumentation

Dialogue

Controversial issue

Democratic citizenship

ABSTRACT

Amid rising polarization, misinformation and disinformation, exclusionary discourse, and global conflict, developing individuals' reasoning skills for contemporary controversial issues constitutes a pressing educational priority, yet evidence on effective approaches remains limited. This study addresses this gap by examining whether a dialogic intervention on a "neutral" topic—based on argumentation and reflection—can develop reasoning skills that transfer to a controversial sociopolitical issue of personal relevance. One hundred twelve sixth graders participated in nine 40-min sessions focused on either US immigration or soda drink consumption to examine near and far transfer. Argumentation skills were assessed through individual essays before and after the intervention on both the intervention and transfer topics. Results showed that both conditions demonstrated gains in argumentation skills, particularly in using evidence, constructing counterarguments, and engaging in two-sided reasoning on both topics. Of particular interest is the far transfer observed from a neutral science topic to a topical, emotionally and belief-laden sociopolitical issue with minimal content overlap. We suggest this achievement is attributable to students' developing meta-level appreciation of the role of evidence and multi-perspectivity in reasoning. Implications for research on developing argumentation skills and the teaching of controversial contemporary topics are discussed.

Democratic societies depend on citizens who can reason effectively on societal issues, which are often controversial and emotionally laden. Yet the educational systems tasked with cultivating this capacity are failing to do so at a time when it is needed most. The heightened political polarization (Pew Research Center, 2023), the rise of far-right populism and exclusionary discourses (Liger et al., 2022; Mudde, 2019), and sustained global conflicts (Institute for Economics and Peace, 2025) that we are witnessing are alarming. Paradoxically, these developments coincide with record highs of formal education: educational attainment has increased nearly fourfold over the past half-century (U.S. Census Bureau, 2017). This suggests that formal education has not succeeded in cultivating the critical thinking skills and democratic competencies required to address contemporary societal challenges. The growing use of social media, especially by young generations, not only reproduces but often intensifies polarized views and creates fertile ground for the spread of misinformation, xenophobia, and ethnonationalist ideologies (Kubin & von Sikorski, 2021).

In this evolving landscape, which reinforces one's viewpoint through social media "echo chambers", the ability to consider multiple

perspectives has become increasingly critical. Such capacity is fundamental for making better judgments and developing deeper understanding through engagement with alternative approaches and viewpoints (Barzilai & Weinstock, 2020). In the absence of gatekeeping, fact-checking, multi-perspectivity, accountability, or sufficient regulatory oversight in social media and online forums, which constitute threats to democracy (Christodoulou & Iordanou, 2021), education's role in supporting individuals to develop their critical thinking skills and dispositions is of paramount importance today. Schools, as microcosms of society, are not immune to these broader socio-political dynamics. On the contrary, they are increasingly confronted with the emotional and ideological spillover of these public discourses (Zembylas, 2021). Yet, evidence-based pedagogical guidelines on how to promote the skills needed to deal with real-life controversial issues are lacking.

Teaching controversial or politically sensitive issues is perceived as risky or divisive (Hess, 2002, 2004; Kerr & Huddelstone, 2015) with teachers reporting that they face pedagogical dilemmas when teaching controversial issues and acknowledging the need to find ways to combine empathy with critical, evidence-based reflection (Zembylas

* Corresponding author at: 12 – 14 University Avenue Pyla, 7080 Larnaka, Cyprus.

E-mail address: KIordanou@uclan.ac.uk (K. Iordanou).

et al., 2020). As a result, scholars and educators face a pressing challenge: How can education support the development of students' critical reasoning to actively engage with controversial issues, rather than passively absorb or reproduce biased and stereotypical beliefs? The answer is not straightforward. Research shows that some efforts aimed at improving critical reasoning on contemporary issues may backfire—for instance, exposure to opposing views can increase polarization (Bail et al., 2018)—or may fail to transfer to real-world contexts (Korteling et al., 2018). In the present work, we examined whether extended engagement in dialogic argumentation and reflection can support the development of reasoning skills that transfer beyond the immediate context to contemporary, controversial real-life societal issues of personal relevance to the students.

1. Development of argumentation skills

At the heart of democratic competencies is individuals' ability to reason well, which involves advanced argumentation skills, particularly the ability to engage in evidence-based reasoning and consider multiple perspectives, to critically evaluate arguments and evidence, rather than making judgments based on unexamined prior beliefs driven by my-side bias (Iordanou et al., 2020; Stanovich & West, 2007). Yet, evidence shows that individuals tend to think mono-perspectively and do not rely on evidence when making judgments (Kuhn & Iordanou, 2022). This is concerning, as failure to engage in critical thinking, such as evaluation of evidence, in favor of reliance on intuitive thinking, makes individuals more prone to conspiracy theories (Binnendyk & Pennycook, 2022; Georgiou et al., 2021), while cognitive inflexibility, that is, failure to switch perspectives, predicts extremist attitudes (Zmigrod et al., 2019).

Mounting evidence from empirical research shows that argumentation skills are amenable to improvement when they receive direct attention (Firetto et al., 2025; Iordanou & Kuhn, 2025; Murphy & Firetto, 2018; Rapanta & Felton, 2022; Shi, 2019). Engagement in purposeful, extensive practice in dialogic argumentation and reflective activities (see Iordanou & Rapanta, 2021 and Kuhn, 2025, for a review) supports the development of students' argumentation skills, particularly the development of the ability to construct two-sided and evidence-based arguments (Iordanou & Constantinou, 2015; Iordanou & Fotiou, 2025; Kuhn et al., 2008). Microgenetic research in argumentation reveals that advanced argumentative skills, such as counterargument, develop gradually through consistent practice. Alongside this development, two forms of understanding co-emerge: meta-strategic understanding, i.e., knowing when and why to use particular strategies, and epistemic understanding, i.e., the disposition to engage in argumentation and one's standards for what constitutes persuasive evidence in belief adoption and revision (Iordanou, 2022; Kuhn et al., 2008). Students progressively develop and internalize an appreciation for and adherence to discourse norms (Iordanou, 2022; Kuhn & Zillmer, 2015; Shi, 2020), including paying attention to others' positions, which gradually helps them to expand their thinking, moving from one, usually my-side, perspective to a multiperspective approach, appreciating different perspectives on an issue (Iordanou & Fotiou, 2025). They expect claims to have reasons to support them, which must stand up to the challenge of strong arguments and evidence that may weaken them. The simultaneous growth of these meta-level insights and procedural argumentation competencies underscores metacognition's pivotal role as the mechanism behind gains in argumentative skills and their transfer to novel contexts (Iordanou, 2022). Yet, although transfer of argumentation gains has been documented across topics, domains and communication modes (see Iordanou & Rapanta, 2021, for a review), our understanding of whether these skills transfer to real-life controversial, affectively laden societal issues of personal relevance remains limited. Controversial issues are defined as those issues that lack consensus and are contested in public and private discourse; such issues are not only intellectually demanding but also charged (Hess, 2002, 2004; Kerr & Huddelstone, 2015). The topics used in previous implementations of

this intervention, both for intervention and for examining transfer effects, were about controversial societal topics, but ones for which participants had no or limited personal or emotional involvement. For example, the topics that have been used were about the social security system (Iordanou et al., 2019), the criteria that a country should employ for immigration (Iordanou & Fotiou, 2025), and climate change (Iordanou & Kuhn, 2025). Previous research suggests that reasoning about personally involving and emotionally laden societal topics poses unique challenges. When individuals hold strong personal stakes, they are more susceptible to my-side bias (Iordanou et al., 2020; Stanovich et al., 2013) and tend to selectively evaluate evidence in ways that confirm their existing beliefs (Iordanou et al., 2025; Kłaczynski, 2000). Emotionally charged topics may also trigger identity-protective reasoning, making it harder for individuals to engage in the kind of decentration that effective argumentation requires (Kahan, 2013). Despite these challenges, very few studies have examined whether argumentation interventions can help individuals overcome these tendencies when discussing issues of personal relevance. Whether the argumentation skills developed through this intervention are robust enough to transfer when reasoning about societal issues where participants hold strong emotions and personal stakes remains an open and important question—one with significant implications for civic education and democratic participation. This study seeks to fill this gap, examining the extent to which argumentation skills can be transferred to the discussion of controversial societal issues with personal stakes.

2. The present study

In the present work, we examined whether a dialogic intervention method, which is based on dialogic argumentation and reflection, on a “neutral” topic can support the development of reasoning skills that are transferable to a controversial real-life societal problem of personal relevance. Specifically, we investigated whether engagement in structured argumentation, embedded in respectful and evidence-based discussion, can foster what Piaget (1932) described as decentration: the capacity to move beyond egocentric thinking by recognizing and considering multiple viewpoints. We were particularly interested in whether this ability, once developed, transfers to controversial topics with personal stakes. Drawing on pre- and post-intervention data, we explored whether students' argumentation skills showed gains on a controversial real-life topic about which they held strong emotions and exhibited my-side bias (Iordanou et al., 2020). We focused particularly on evidence use and acknowledgment of alternative positions, as these lie at the heart of advanced argumentation skills and are fundamental for democratic citizenship.

The controversial societal issue with personal stakes we used was the topic of Turkish settlers. In the Cypriot context, the term *Turkish settlers* refers to individuals from mainland Turkey who migrated to the northern part of Cyprus following the island's division in 1974. Their presence remains a deeply contested political and moral issue, linked to questions of sovereignty, demographic change, settler colonialism, and the legitimacy of the post-1974 political order (Moudouros, 2025; Zembylas, forthcoming). Among Greek Cypriots, discussions about Turkish settlers often evoke strong emotions tied to displacement, occupation, and national identity, making the topic highly sensitive in educational settings. Existing research in the Greek-Cypriot community over the last 25 years suggests that children at the end of primary school are already exposed to and often internalize dominant narratives about the 1974 conflict and its aftermath. For example, studies with Greek-Cypriot primary school students have shown that by late primary grades many students are familiar with key historical events and reproduce widely circulating societal discourses about Turkish Cypriots and Turkish settlers, often drawing on narratives encountered in family conversations, media, and school commemorative practices (e.g., Christou, 2007; Papadakis, 2008; Spyrou, 2006; Zembylas, 2013; Zembylas & Michaelidou, 2013). This research has documented that

Greek-Cypriot students in upper primary school often encounter emotionally charged representations of the conflict through school commemorations, textbooks, and classroom discussions, which contribute to shaping early affective orientations toward the “other” (e.g., Zembylas et al., 2016; Zembylas & Kambani, 2012; Zembylas & Loukaidis, 2021).

This study makes two key contributions to existing research. First, our work adds to the growing body of scholarship that links argumentation to democratic education (e.g., Joshi, 2016; Kuhn, 2025; Moshman, 2020; Rapanta & Iordanou, 2023; Weinstock, 2005), highlighting its potential not only as a critical thinking skill but as a mode of civic formation. Argumentation, when designed as a dialogic and participatory practice, offers a powerful antidote to the echo chambers and tribalism that characterize much of today's public discourse (Barberá et al., 2015). Moreover, it can help students resist simplistic narratives, including misinformation and disinformation, conspiracy theories, and stereotypes, by encouraging deeper engagement with complexity, ambiguity, and difference. Second, our work extends research on the development of argumentation skills by examining far transfer to a current, controversial, and affectively laden topic of personal relevance to the students (Iordanou et al., 2025). Transfer research distinguishes between near transfer, where the training and transfer contexts share surface-level features, and far transfer, where the contexts differ substantially in content and structure (Barnett & Ceci, 2002). Demonstrating far transfer provides particularly strong evidence that learners have developed generalizable skills rather than context-specific knowledge. Our study examines both types of transfer.

To address our research question, “Can argumentation skills developed through an intervention based on dialogic argumentation and reflection on a neutral topic transfer to a controversial real-life problem of personal relevance?”, we had 6th graders engage in the ‘Argue-with-Me’ curriculum (Iordanou & Rapanta, 2021) and examined their argumentation skills before and after their engagement in the intervention. We hypothesized that “neutral” topics, which do not carry the burden of heightened emotional reactivity that impairs complex cognitive processing (Mason et al., 2017; Mason et al., 2019), could provide favorable conditions for developing reasoning skills, and that once developed, these skills would transfer to more emotionally charged controversial issues. Therefore, we chose to have students first develop their argumentation skills on an emotionally neutral topic before examining transfer to an emotionally laden personal issue. Research on teaching controversial issues suggests this tactic to address emotionally difficult issues, without implying that transfer across emotionally distinct contexts is automatic (Hess, 2009). From this perspective, then, emotionally neutral topics provide an initial learning environment in which students can acquire and stabilize core argumentative practices before testing whether these practices remain available under conditions of greater emotional and identity-related challenge. Once these skills are sufficiently developed and internalized as part of one's reasoning repertoire, they may become robust enough to be deployed even in emotionally charged contexts. In other words, the neutral topic serves as a scaffolding context in which students can focus on developing the structural and metacognitive aspects of argumentation without the interference of emotional and identity-related biases. The question then becomes whether these skills, once consolidated, can withstand the added demands of reasoning about a personally relevant controversial issue. To examine this, we designed two experimental conditions. We had one group of students discuss the issue of undocumented immigrants in the US—a topic with similarities to the transfer topic of Turkish settlers in Cyprus, yet ‘neutral’ for students in Cyprus as it involved another country far away from their own—constituting a near-transfer condition. Furthermore, to rule out the possibility that any gains in participants' argumentation skills on the transfer topic could be attributed to familiarity with the immigration topic, we included a far-transfer condition: another group that engaged in argumentation on an unrelated topic, the consumption of soft drinks. This condition would enable us to

document far transfer with greater certainty, demonstrating transfer from a scientific topic to a sociopolitical and emotionally laden topic of personal stake. For both topics, students gradually received background information in the form of question-and-answer cards, which they could use in their dialogs to develop evidence-based arguments and counter-arguments (Iordanou & Kuhn, 2025).

3. Methods

3.1. Participants

A total of 116 6th grade students were invited to participate in the study, representing the entire sixth-grade population of three public schools in Cyprus. Two of the schools each had two classes, and one school had three classes, resulting in a total of seven classes. For School A, there were 42 students in total, with 22 students in Class 1 and 20 students in Class 2. For School B, there were 34 students, with 15 students in Class 1 and 19 students in Class 2. For School C, there were a total of 40 students across three classes: 17 students in Class 1, 7 students in Class 2, and 16 students in Class 3. This study was approved by the Cyprus National Bioethics Committee (Approval No.: EEBK EII 2022.01.02, Date: December 15, 2022). All procedures were conducted in compliance with relevant national laws and institutional guidelines. Written informed consent was obtained from legal guardians, and verbal assent was obtained from all child participants. The privacy rights of participants were protected throughout the study. Of the 116 students invited, 112 (64 girls, 57.1%; 47 boys, 42.0%; 1 non-binary, 0.9%), aged 11 to 12 years ($M = 11.43$, $SD = 0.51$), provided informed consent for their data to be included in the analysis. The schools that participants attended served mostly middle-class families and within an average range of ability and academic achievement. The participants were mostly Greek-Cypriots (90%), with a minority being a mixture of other ethnicities (10%).

3.2. Design

Seven comparable 6th grade classes were randomly assigned to one of two conditions based on the proximity of the intervention topic to the transfer topic: the near-transfer condition ($n = 59$), which engaged with the thematically related topic of undocumented immigrants in the US, and the far-transfer condition, which engaged with the unrelated topic of soft drinks consumption ($n = 53$). Participants in the far-transfer condition consisted of three classes, one for each school. Participants in the near-transfer condition consisted of four classes: one from School A, one from School B and two from School C.

3.3. Procedure

The intervention consisted of nine 40-min sessions plus two sessions for assessment, one for initial and another one for final assessment. All participants engaged in the same activities, with only the topic being different between the two conditions for Phase I (sessions 1 to 8). For Phase II (session 9) both conditions engaged with the same topic (see Table 1). Participants in the near-transfer condition were asked in Phase I to discuss the topic of illegal immigrants in the USA, and specifically discuss the following question: “What do you think should happen to the children of undocumented immigrants who were brought to the US as children and are now living there illegally?” Participants in the far-transfer condition were asked to discuss in Phase I whether they think that people who like soft drinks, have a balanced diet, and are relatively healthy can consume soft drinks in moderation or whether they think that soft drinks have no benefit and are best to be avoided. In Phase II, students of both conditions were asked to discuss a real-life sociopolitical topic in their country for 1 session (session 9): whether Turkish settlers who now live in the northern (occupied) part of Cyprus should continue to live in these houses, or should be given other houses by the

Table 1
Overview of Study Phases and Sessions.

Phase	Session	Activity	Description
–	Initial	Initial Assessment – Intervention & Transfer Topics	Students wrote an essay on the intervention topic and a second essay on the transfer topic.
I	1–4	Dialogs & Reflection	Same-side pairs engaged in four 20-min electronic dialogs with opposing-position pairs. From Session 2 onward, students also completed reflective activities using own-side and other-side reflection sheets. Q&A cards accumulated across sessions.
I	5	Preparation of Showdown	Groups of 5–6 same-side students prepared arguments for the showdown, divided into “own argument” and “other argument” specialists.
I	6	Showdown	Two teams per side debated via IM software, with roles reversed halfway through.
I	7	Feedback	Researchers provided feedback using argument maps and a point system based on the showdown dialogue.
I	8	Final Assessment – Intervention Topic	Students wrote an essay on the intervention topic and were introduced to the transfer topic.
II	9	Dialogue	Same-side pairs engaged in two 20-min electronic dialogs on the transfer topic with opposing-position pairs, supported by Q&A cards.
II	10	Final Assessment – Transfer Topic	Students wrote an essay on the transfer topic.

government and move out from the houses that originally belonged to Greek Cypriots.

3.4. Initial assessment

Participants were asked to write a brief essay on their intervention topic. Those in the near-transfer condition took a position on whether undocumented immigrants should stay or leave the US. Those in the far-transfer condition took a position on whether healthy individuals who maintain balanced diets can consume soft drinks or should avoid them entirely. Both conditions included an undecided option. For the Turkish-Settlers topic, students in both conditions were asked to take a position on whether Turkish settlers should stay in the houses they currently reside in, move to other houses or whether one was undecided. The prompt for all the topics was to “Write the argument you would make to convince someone who does not agree with your choice. Write the best argument you can”.

Table 2
Means, Standard Deviations, and 95% Confidence Intervals for Idea Units, Evidence, and Types of Evidence Function on the Intervention Topic, by Condition and Time.

	Far-Transfer Condition (n = 53)				Near-Transfer Condition (n = 59)			
	Pre-Test		Post-Test		Pre-Test		Post-Test	
Idea Units	M (SD)	95% CI	M (SD)	95% CI	M (SD)	95% CI	M (SD)	95% CI
Evidence	2.04 (1.01)	[1.75, 2.34]	4.72 (1.99)	[4.17, 5.27]	2.36 (0.98)	[2.10, 2.61]	5.91 (2.87)	[5.16, 6.67]
Functional Evidence	1.13 (1.18)	[0.78, 1.47]	4.30 (2.00)	[3.75, 4.85]	2.00 (0.95)	[1.75, 2.25]	5.38 (2.69)	[4.67, 6.09]
Support_Own	1.06 (1.12)	[0.74, 1.39]	4.15 (1.90)	[3.63, 4.67]	1.92 (0.95)	[1.67, 2.16]	5.19 (2.63)	[4.50, 5.88]
Weaken_Other	0.83 (1.02)	[0.54, 1.13]	3.06 (1.49)	[2.65, 3.47]	1.27 (1.06)	[0.99, 1.55]	2.48 (2.16)	[1.91, 3.05]
Support_Other	0.10 (0.31)	[0.02, 0.19]	0.83 (1.16)	[0.51, 1.15]	0.63 (0.91)	[0.39, 0.86]	2.22 (1.84)	[1.74, 2.71]
Weaken_Own	0.06 (0.25)	[−0.01, 0.13]	0.26 (0.96)	[−0.00, 0.53]	0.02 (0.13)	[−0.02, 0.05]	0.41 (0.86)	[0.19, 0.64]
Personal	0.00 (0.00)	[0.00, 0.00]	0.02 (0.14)	[−0.02, 0.06]	0.00 (0.00)	[0.00, 0.00]	0.05 (0.29)	[−0.03, 0.13]
QAEvidence	0.44 (0.71)	[0.23, 0.64]	0.64 (0.88)	[0.40, 0.88]	0.71 (0.87)	[0.49, 0.94]	0.86 (1.16)	[0.56, 1.17]
However	0.63 (0.67)	[0.43, 0.82]	3.38 (1.94)	[2.84, 3.91]	1.20 (1.01)	[0.94, 1.47]	4.28 (2.65)	[3.58, 4.97]
	0.19 (0.39)	[0.07, 0.30]	0.21 (0.53)	[0.06, 0.35]	0.27 (0.49)	[0.15, 0.40]	0.38 (0.75)	[0.18, 0.58]

3.4.1. Phase I: intervention topic

3.4.1.1. Phase I. Dialogs & reflection (sessions 1–4). Same-side pairs were formed based on students' positions on the Phase I intervention topic. Each pair engaged in four 20-min electronic dialogs, each with a different opposing-position pair. Before each dialogue, pairs received an information bank of four Q&A cards containing topic-related questions and brief answers. Cards accumulated across sessions (16 total) and were balanced to support both positions. For example, soda cards included “Does soda have nutritional value? Soda lacks essential nutrients like vitamins, minerals, antioxidants, or fiber. Soda is said to have ‘empty’ calories because it mainly contains sugar.” Immigration cards included “Do undocumented immigrants commit more crimes than legal immigrants and US-born citizens? A 2020 study found that undocumented immigrants are involved in fewer crimes compared to legal immigrants and US-born citizens.”

In sessions 2–4, students engaged in a dialogic activity for the first 20 min and in a reflective activity (another 20 mins) using two sheets. The own-side reflection sheet asked students to reflect on the strength of their counterarguments against the opposing side's position. The other-side reflection sheet asked students to reflect on their rebuttals to counterarguments against their own position. In both cases, students were asked to note improved counterarguments or rebuttals using evidence from the Q&A cards.

3.4.1.2. Phase I. Preparation of showdown (session 5). In session 5, 5–6 same-side students prepared for the showdown that would take place in session 6. Having the reflection sheets prepared in previous sessions available, they were asked to decide what arguments to use in the debate. The participants who had been working together as a pair for dialogue and reflection sessions were separated and assigned to two different preparation teams: the “own argument” specialists, working with own-argument reflection sheets, and the “other argument” specialists, working with other-argument reflection sheets. Students used different colored cards for arguments, counterarguments, and rebuttals. An adult coach facilitated these discussions by providing instructions and prompting students to evaluate the strength of their arguments, counterarguments, and rebuttals. The coach also encouraged students to consider how these components could be strengthened before recording them on the cards.

3.4.1.3. Phase I. Showdown (session 6). Participants on each side were divided into two teams (A and B) consisting of 5–6 members. The own-argument and other-argument specialists were represented equally on Team A and Team B. Participants on each side were seated in different rooms so that communication between the two sides took place through IM software, with the dialogue projected on a wall screen. All participants from each side were present in the room, with half actively participating and the other half observing and offering suggestions if they wished. Halfway through, the two groups reversed roles. The

dialogue was projected onto a wall screen in each room. In each room team A debated for the first half of the showdown, and team B did so for the other half. The showdown was presented to students as the end-goal of the intervention, which was framed as a game, drawing on an analogy to athletes preparing for a major competition, to sustain their engagement and motivation throughout. Importantly, although framed as a competition, the showdown was judged on argumentation quality, the use of evidence and engagement with the opposing side, rather than on which position was more persuasive or correct (as reflected in the point system described below).

3.4.1.4. Phase I. Feedback (session 7). The researchers provided feedback to the students, in the form of an argument map that was prepared based on the electronic dialogue produced in the showdown session. Each side's statements were represented in a different column and connected by lines to show their interrelation. Different colors were used to label statements as effective or ineffective argumentative moves, highlighting as effective the use of evidence to support claims and addressing the opponent's position, and as ineffective solely focusing on one's own position without engaging with the other side. A point system was also applied, providing points for argumentative moves addressing the other side's arguments (counter-arguments or rebuttals) and for the use of evidence. Note that the point system was based on argumentation quality rather than evaluating position. The feedback session thus served as the culmination of the Phase I intervention, using rich dialogic material from the showdown for discussing features of advanced argumentation and reflecting on students' own skills.

3.4.1.5. Final assessment – intervention topic (session 8). Students' argument skills on the intervention topic were assessed using the same instrument that was used at initial assessment. During Session 8, students were also presented with the scenario for the transfer topic (Turkish Settlers) and asked to take whichever position they preferred (stay in the houses they currently reside in, move to other houses, or remain undecided) and to write an argument justifying it, as an introduction to the transfer topic that would follow in Phase II.

3.4.2. Phase II: transfer topic

3.4.2.1. Phase II. Dialogue (session 9). Pairs of same-side positions were formed based on students' position on the transfer topic. Students' positions were roughly evenly distributed between the two sides, with undecided ones assigned to the side with fewer students—the “stay” position—to achieve balance between the two groups. Pairs engaged in a 20 min-electronic dialogue with another pair in their classroom who supported the opposing position. Students participated in 2 20 min dialogs with different pairs for each dialogue. Students were also provided with four Q&A cards with relevant pieces of knowledge that they could use in this topic. Some examples of Q&A used were “*Who is the legal owner of the houses that Turkish Settlers currently live in according to international conventions? According to the resolutions of the United Nations Security Council and General Assembly, the Greek Cypriot refugees are the legal owners of those houses*”, and “*Would a potential relocation of the Turkish settlers to houses other than those they currently live in and which were given to them by Turkey be considered unfair? Potentially, yes. That would come in opposition with what Turkey agreed with them in order to encourage them to abandon their homes in their home country and come to the northern (occupied) part of Cyprus*”.

3.4.2.2. Phase II. Final assessment - transfer topic (session 10). Students' argument skills were assessed on the transfer topic, using the same instrument that was used at initial assessment.

3.5. Coding

Students' essays on both the intervention and transfer topics were segmented into idea units by two research assistants, both PhD holders, who had received training from a senior researcher with extensive coding experience (the first author). Firstly, idea units were coded based on whether the piece of evidence was connected to a specific claim or to any broader argument being made. If there was no connection, then the unit was coded as non-functional and was not coded further. If the unit included a piece of evidence which was connected with the claim, for example supporting or weakening it, then it was coded a functional unit. Each functional unit was further categorized according to the type of function it served. The coding system used is the one reported by Kuhn et al. (2016) and by others, based on the type of function the unit serves. Four functions were identified: to support one's own position (*Support my own*); to critique and thereby weaken the opposing position (*Weaken other*); to acknowledge strengths of the opposing position (*Support other*) and to acknowledge weaknesses of one's own position (*Weaken my own*). Plenty of examples of how the essays were coded are provided in the Supplementary Materials of this paper, but for illustration purposes an example of each is also provided here. In an essay where the overall position was that the Turkish Settlers should leave from the houses they are currently residing in, the following idea unit was coded as *support my own*: “The Turkish settlers are not the legal occupants of those houses”, while the following complex sentence comprises two idea units: “Many people believe that the Turkish settlers should stay because they have lived in those houses for almost fifty years (*support other*), but that shouldn't matter because those houses belong by law to Greek Cypriots (*weakening other*)”. Idea units functioning to weaken one's own position were rare in the data. An illustrative example comes from an essay on the topic of illegal immigrants in the USA written during the post-test phase. The pupil's overall stance was that immigrants should be allowed to stay in the country. At the end of the essay, the pupil labeled their final two idea units as counterarguments. One of these stated that remaining in the USA as an illegal immigrant sometimes means that children must take on dangerous jobs that could cost them their lives, essentially weakening their own position in favor of immigrants staying in the USA.

Furthermore, we were interested in examining whether students acknowledged both positions, i.e., their own (Greek-Cypriot) and the others' (Turkish settlers'), as well as whether they constructed “However” arguments. “However” arguments are defined as pairs of adjacent statements that support opposing conclusions (e.g., Support-other + Support-own, or Support-other + Weaken-other), explicitly connected by conjunctions like “however,” “but,” or “although,” demonstrating the arguer's ability to coordinate two contrasting claims simultaneously. The complex sentence provided at the end of the previous paragraph is an example of that. Finally, we examined if there were any meta-statements made about the topic which would include statements that illustrate that the “writer is thinking more deeply about the topic than simply stating his or her own views, for example by noting that the topic is a difficult one or that people have many different views on it” (Kuhn & Modrek, 2021, p. 1353). However, no meta-statements were found in the dataset arguably due to the young age of the participants.

Twenty percent of the data from both the main intervention topic and the transfer topic were jointly coded by both raters as part of their training to establish agreement in their coding approach. The inter-rater reliability was assessed using 20% of the remaining data for coding argumentation skills across both main intervention and transfer topic codes. Cohen's Kappa demonstrated almost perfect agreement, $K = 0.85$, $p < .001$, $z = 22$, suggesting a high level of consistency in coding practices among the two raters. Any disagreements in coding were resolved through discussion among all three coders, including the senior researcher, and final decisions were reached through dialogue and consensus. All the transfer topic data were coded by two coders. The remaining dataset was divided between the two trained coders who coded them independently, after establishing high inter-rater

agreement.

We examined whether participants in the two conditions—near- and far-transfer—showed gains in their argumentation skills on their respective intervention topic and whether these gains transferred to a controversial sociopolitical issue with personal stakes for students in Cyprus, i.e., the Turkish settlers topic. We compared performance from pre- to post-intervention on evidence use, the functionality of evidence used (supporting own position, weakening other position, supporting other position, and weakening own position), construction of “However” arguments, and acknowledgment of alternative positions, as these lie at the heart of advanced argumentation skills. Separate Poisson GLMMs were fitted for each variable to examine changes from pre- to post-intervention, controlling for essay length and the nesting of students within classrooms.

4. Results

4.1. Transfer topic

Data were screened for outliers using the interquartile range method (IQR). No extreme outliers were identified. All cases were retained for subsequent analyses.

4.2. Equivalence of experimental conditions at baseline

Independent samples *t*-tests were conducted to examine baseline differences between the two conditions across all measured variables, showing that the two groups were comparable at initial assessment across all measured constructs. No statistically significant differences were identified at pre-test between the conditions on any of the assessed variables: IdeaUnitsPre ($t(110) = 1.58, p = .117, d = 0.30$), Evidence_Pre ($t(110) = 1.53, p = .130, d = 0.29$), Evid_Fun_Pre ($t(110) = 1.32, p = .190, d = 0.25$), SuppOwn_Pre ($t(110) = 0.75, p = .455, d = 0.14$), SupportOther_Pre ($t(110) = 0.59, p = .554, d = 0.11$), PersonalEvid_Pre ($t(110) = 1.00, p = .321, d = 0.19$), QA_Ev_Pre ($t(110) = 0.93, p = .357, d = 0.18$), and However_Pre ($t(110) = 1.25, p = .214, d = 0.24$). For WeakenOther_Pre, due to violations of normality assumptions, a Mann-Whitney *U* test was conducted instead of an independent samples *t*-test, which revealed no significant difference between Group 1 and Group 2, $U = 1498.50, z = -0.56, p = .577, r = 0.05$.

4.3. Intervention effects over time on the us immigration topic

Before examining transfer effects, we first assessed whether participants in the near-transfer condition, who engaged with the thematically related topic of undocumented immigrants in the US, showed gains in their argumentation skills on this intervention topic.

4.3.1. Idea units

A GLMM with a Poisson distribution and log link function examining changes in idea unit production over time, controlling for class, showed a significant main effect of Time, $F(1,113) = 121.16, p < .001$. Students significantly increased their idea unit production from pre-test ($M = 2.36, SE = 0.13, 95\% CI [2.12, 2.62]$) to post-test ($M = 5.91, SE = 0.38, 95\% CI [5.21, 6.71]$).

4.4. Argumentation skills

Poisson GLMMs were fitted with the natural log link to examine changes in evidence used, functional evidence used, the type of function of evidence, the source of evidence and the construction of “However” arguments. To examine whether changes in argument skills reflected genuine improvements in argumentation quality rather than increased essay elaboration and to account for the nested structure of classrooms, idea units and class were included as fixed effects.

4.4.1. Evidence used

The model revealed that Time was a significant predictor of evidence use, $F(1,112) = 44.10, p < .001$, even after controlling for essay length (idea units) and classroom membership. Estimated marginal means indicated that participants produced significantly more evidence at post-test ($M = 3.82, SE = 0.12, 95\% CI [3.59, 4.06]$) compared to pre-test ($M = 2.52, SE = 0.13, 95\% CI [2.27, 2.79]$).

4.4.2. Functional evidence used

The model revealed that Time was a significant predictor of functional evidence use, $F(1, 112) = 36.96, p < .001$, even after controlling for essay length (idea units) and classroom membership. Estimated marginal means indicated that participants produced significantly more functional evidence at post-test ($M = 3.67, SE = 0.13, 95\% CI [3.42, 3.94]$) compared to pre-test ($M = 2.41, SE = 0.13, 95\% CI [2.16, 2.69]$).

4.4.3. Support own (sup_own)

Students showed comparable performance in the usage of evidence to support their own position at pre-test ($M = 1.58, SE = 0.18, 95\% CI [1.26, 1.98]$) and post-test ($M = 1.82, SE = 0.21, 95\% CI [1.44, 2.29]$), $F(1, 112) = 0.64, p = .426$.

4.4.4. Weaken other (weak_other)

Significant change was observed in the usage of evidence which functioned to weaken other's position, $F(1, 112) = 5.92, p = .017$. Students improved from pre-test ($M = 0.72, SE = 0.15, 95\% CI [0.48, 1.07]$) to post-test ($M = 1.29, SE = 0.17, 95\% CI [1.00, 1.66]$).

4.4.5. Support other (sup_oth)

Students also demonstrated significant improvement in using evidence to support other's position, $F(1, 112) = 5.37, p = .022$. Performance increased from pre-test ($M = 0.02, SE = 0.02, 95\% CI [0.00, 0.12]$) to post-test ($M = 0.18, SE = 0.08, 95\% CI [0.07, 0.45]$).

4.4.6. Weaken own (weak_own)

In contrast, no significant time effect was found for evidence used to weaken one's own position, $F(1, 112) = 2.33, p = .130$. This type of evidence use was virtually absent at pre-test ($M \approx 0.00, SE \approx 0.00$) and showed only negligible use at post-test ($M \approx 0.00, SE \approx 0.00$).

4.4.7. Source of evidence

4.4.7.1. Personal evidence (Pers_Ev). No significant time effect was observed for personal evidence use, $F(1, 112) = 0.46, p = .499$. Students showed similar levels of performance from pre-test ($M = 0.80, SE = 0.15, 95\% CI [0.56, 1.16]$) to post-test ($M = 0.65, SE = 0.15, 95\% CI [0.41, 1.01]$).

4.4.7.2. Q&A evidence (QA_Ev). Students demonstrated significant improvement in using evidence from Q&A cards, $F(1, 112) = 27.18, p < .001$.² Usage increased substantially from pre-test ($M = 1.55, SE = 0.16, 95\% CI [1.26, 1.90]$) to post-test ($M = 3.00, SE = 0.19, 95\% CI [2.66, 3.39]$).

4.4.8. However arguments

No significant time effect was found for “however” usage, $F(1, 112) = 0.86, p = .355$. This type of usage remained minimal from pre-test ($M = 0.33, SE = 0.09, 95\% CI [0.19, 0.57]$) to post-test ($M = 0.22, SE = 0.07, 95\% CI [0.12, 0.40]$).

¹ Estimated marginal means reported here are derived from the GLMM and are adjusted for class and text length. Raw descriptive means and standard deviations are presented in Table 2.

² Both effects were evaluated against a Bonferroni-corrected significance threshold of $\alpha = 0.025 (0.05/2)$ to account for the two dependent variables.

4.4.9. Reference to both positions

A generalized linear mixed model (GLMM) with a binomial distribution and logit link function was fitted to examine whether the likelihood of referring to both positions changed over time (pre vs. post). The overall model, controlling for class, was statistically significant, $F(3,113) = 6.67, p < .001$. Time was a significant predictor, $F(1, 113) = 19.04, p < .001$. Estimated marginal means indicated that the probability of referencing both positions increased significantly from pre-test ($M = 0.21, SE = 0.06, 95\% \text{ CI } [0.12, 0.34]$) to post-test ($M = 0.65, SE = 0.07, 95\% \text{ CI } [0.51, 0.77]$). This increase was statistically significant, with a pairwise contrast estimate of 0.44 ($SE = 0.09, t(113) = 5.12, p < .001, 95\% \text{ CI } [0.27, 0.62]$).

4.5. Intervention effects over time on the soft drinks topic

As with the near-transfer condition, we examined whether participants who engaged with the soft drinks topic showed improvement in their argumentation skills from pre- to post-intervention on the intervention topic.

4.5.1. Idea units

A GLMM with a Poisson distribution and log link function was used to examine changes in idea unit production over time, controlling for Class. The overall model was significant, $F(3,100) = 31.64, p < .001$. There was a significant main effect of Time, $F(1,100) = 89.12, p < .001$, with students significantly increasing their idea unit production from pre-test ($M = 2.10, SE = 0.14, 95\% \text{ CI } [1.84, 2.40]$) to post-test ($M = 4.79, SE = 0.27, 95\% \text{ CI } [4.28, 5.35]$).

4.6. Argumentation skills

Poisson GLMMs were fitted with the natural log link to examine changes in functional evidence used, the type of function of evidence, the source of evidence, and the construction of “However” arguments, controlling for idea units and class, which were included as fixed effects.

4.6.1. Evidence used

The model revealed that Time was a significant predictor of evidence use, $F(1, 99) = 28.30, p < .001$, even after controlling for essay length (idea units) and classroom membership. Estimated marginal means indicated that participants produced significantly more evidence at post-test ($M = 3.06, SE = 0.12, 95\% \text{ CI } [2.83, 3.31]$) compared to pre-test ($M = 1.49, SE = 0.19, 95\% \text{ CI } [1.16, 1.92]$).

4.6.2. Functional evidence use

Regarding functional use of evidence, in particular, a GLMM showed a significant time effect, $F(1, 99) = 29.52, p < .001$. Students substantially improved from pre-test ($M = 1.41, SE = 0.18, 95\% \text{ CI } [1.09, 1.82]$) to post-test ($M = 3.01, SE = 0.13, 95\% \text{ CI } [2.76, 3.28]$).

4.6.3. Support own (sup_own)

Regarding the functionality of evidence, a significant effect of Time was found for evidence used to support one's own position, $F(1, 99) = 23.13, p < .001$. Adjusted estimated marginal means increased substantially from pre-test ($M = 1.01, SE = 0.16, 95\% \text{ CI } [0.73, 1.38]$) to post-test ($M = 2.50, SE = 0.21, 95\% \text{ CI } [2.12, 2.94]$).

4.6.4. Weaken other (weak_other)

For the more challenging usage of evidence aiming to weaken other positions, a significant change was observed, $F(1, 99) = 4.90, p = .029$. Adjusted estimated marginal means increased from pre-test ($M = 0.14, SE = 0.06, 95\% \text{ CI } [0.06, 0.32]$) to post-test ($M = 0.43, SE = 0.11, 95\% \text{ CI } [0.27, 0.71]$).

4.6.5. Support other (sup_oth)

Regarding using evidence with the function to support other

positions, no significant effect of Time was found, $F(1, 99) = 0.00, p = .999$. Estimated marginal means were near zero at both time points (Time 1: $M = 0.00, SE = 0.55$; Time 2: $M = 0.00, SE = 0.14$).

4.6.6. Weaken own (weak_own)

Regarding evidence used to weaken one's own position, the model experienced estimation difficulties due to extremely low base rates across conditions, rendering the Time and Class effects non-estimable (0 degrees of freedom). This type of evidence use remained virtually absent at both time points.

4.6.7. Source of evidence

4.6.7.1. Personal evidence. Regarding the source of evidence (personal or from QA cards) the students showed no statistically significant change in personal evidence usage between initial and final assessment, $F(1, 99) = 1.38, p = .243$. Students showed moderate usage of personal evidence at pre-test ($M = 0.61, SE = 0.13, 95\% \text{ CI } [0.40, 0.92]$) and post-test ($M = 0.38, SE = 0.11, 95\% \text{ CI } [0.21, 0.68]$), with no statistically significant difference between time points.

4.6.7.2. QA evidence (qa_evi). In contrast, students showed a statistically significant increase in QA card evidence usage between initial and final assessment, with a significant main effect of Time, $F(1, 99) = 37.86, p < .001$. Students used considerably more QA card evidence at post-test ($M = 2.32, SE = 0.20, 95\% \text{ CI } [1.96, 2.74]$) compared to pre-test ($M = 0.81, SE = 0.11, 95\% \text{ CI } [0.62, 1.07]$). This significant increase suggests that, following the intervention, students shifted toward drawing more heavily on evidence from the QA cards rather than relying on personal experience alone.

4.6.8. However arguments

No significant change was observed for the usage of “however” arguments, $F(1, 96) = 0.00, p = 1.000$. Students showed minimal usage at both pre-test ($M = 0.18, SE = 0.16, 95\% \text{ CI } [., 0.00]$) and post-test ($M = 0.13, SE = 0.06, 95\% \text{ CI } [0.05, 0.33]$), with no statistically significant change between time points.

4.6.9. Reference to both positions

A GLMM with a binomial distribution and logit link function was fitted to examine whether the likelihood of referring to both positions changed over time (pre vs. post). The overall model, controlling for class, was statistically significant, $F(3, 98) = 3.47, p = .019$. Time was a significant predictor, $F(1, 98) = 6.55, p = .012$. Estimated marginal means indicated that the probability of referencing both positions increased significantly from pre-test ($M = 0.66, SE = 0.07, 95\% \text{ CI } [0.50, 0.78]$) to post-test ($M = 0.89, SE = 0.05, 95\% \text{ CI } [0.76, 0.95]$), with a pairwise contrast estimate of 0.23 ($SE = 0.09, t(98) = 2.70, p = .008, 95\% \text{ CI } [0.06, 0.40]$).

4.7. Intervention effects over time on the transfer topic

After documenting gains in argumentation skills on the intervention topics, we examined whether these skills transferred to the Turkish settlers topic—a controversial sociopolitical issue with personal stakes for students in Cyprus. We compared performance from pre- to post-intervention for both the near-transfer and far-transfer conditions.

4.7.1. Idea units

To assess the effects of Group and Time on idea units in students' essays, a generalized linear mixed model with Poisson distribution was used, accounting for repeated measures within students and controlling for class membership. The overall model was statistically significant, $F(8, 215) = 15.93, p < .001$. The fixed effect of Group was not statistically significant ($p = .905$), nor was the Group $\times$ Time interaction, $F(1, 215)$

= 0.34, $p = .563$. The fixed effect of Time was statistically significant, $F(1, 215) = 97.33$, $p < .001$. Estimated marginal means for idea units increased from 2.41³ ($SE = 0.25$) to 5.11 ($SE = 0.42$) in the Soda condition and from 2.48 ($SE = 0.22$) to 4.84 ($SE = 0.35$) in the US condition. Both groups showed significant improvements from pre- to post-test (Soda: $p < .001$; US: $p < .001$), representing substantial increases in the number of idea units students included in their argumentative essays following the intervention.

4.7.2. Evidence used

To assess the effects of Group and Time on the use of evidence relative to total idea units, a generalized linear mixed model with Poisson distribution was used, with idea units included as a covariate, accounting for repeated measures within students and controlling for class membership. The overall model was statistically significant, $F(9, 214) = 77.24$, $p < .001$. Time was a significant predictor of evidence use, $F(1, 214) = 17.88$, $p < .001$, while neither the fixed effect of Group ($p = .154$), Class, $F(5, 214) = 1.97$, $p = .084$, nor the Group $\times$ Time interaction, $F(1, 214) = 2.03$, $p = .156$, were statistically significant. Estimated marginal means indicated that participants produced significantly more evidence at post-test ($M = 3.22$, $SE = 0.12$, 95% CI [3.00, 3.46]) compared to pre-test ($M = 2.55$, $SE = 0.10$, 95% CI [2.35, 2.76]). Pairwise contrasts within each group revealed that both the US condition (pre: $M = 2.72$, $SE = 0.15$; post: $M = 3.22$, $SE = 0.12$; $p = .014$) and the Soda condition (pre: $M = 2.38$, $SE = 0.16$; post: $M = 3.22$, $SE = 0.12$; $p < .001$) showed significant increases in evidence use over time. The non-significant interaction suggests that the improvement in evidence use from pre- to post-test did not differ significantly between the two topic conditions.

4.7.3. Functional evidence use

To assess the effects of Group and Time on the use of functional evidence relative to total idea units, a generalized linear mixed model with Poisson distribution was used, with the natural log of idea units included as an offset, accounting for repeated measures within students and controlling for class membership. The overall model was not statistically significant, $F(8, 213) = 0.51$, $p = .849$. Neither the fixed effect of Group ($p = .351$), Time, $F(1, 213) = 0.01$, $p = .933$, nor the Group $\times$ Time interaction, $F(1, 213) = 0.71$, $p = .401$, were statistically significant. Estimated marginal means for the rate of functional evidence use remained stable from pre-test ($EMM = 0.84$, $SE = 0.03$) to post-test ($EMM = 0.84$, $SE = 0.02$), and neither the US ($p = .569$) nor the Soda condition ($p = .538$) showed significant changes over time. These results suggest that while students produced significantly more idea units following the intervention, the proportion of those idea units that contained functional evidence did not change.

4.7.3.1. Support-own use. To assess the effects of Group and Time on the rate of Support-Own evidence use relative to total idea units, a generalized linear mixed model with Poisson distribution was used, with the natural log of idea units included as an offset, accounting for repeated measures within students and controlling for class membership. The overall model was statistically significant, $F(8, 213) = 4.48$, $p < .001$. Neither the fixed effect of Group ($p = .102$) nor the Group $\times$ Time interaction, $F(1, 213) = 0.09$, $p = .766$, were statistically significant. The fixed effect of Time was statistically significant, $F(1, 213) = 21.91$, $p < .001$. Although raw counts of Support-Own evidence increased from pre- to post-test, the estimated marginal mean rate of Support-Own evidence use relative to total idea units significantly decreased from pre-test ($EMM = 0.69$, $SE = 0.03$) to post-test ($EMM = 0.50$, $SE = 0.03$).

³ Estimated marginal means reported here are derived from the GLMM and are adjusted for class and text length. Raw descriptive means and standard deviations are presented in Table 3.

4.7.4. Weaken-other

For the rate of Weaken-Other evidence use, a generalized linear mixed model with Poisson distribution and the natural log of idea units as an offset was used, accounting for repeated measures within students and controlling for class membership. The overall model was statistically significant, $F(8, 213) = 4.27$, $p < .001$. Neither the fixed effect of Group ($p = .210$) nor the Group $\times$ Time interaction, $F(1, 213) = 0.00$, $p = .986$, were statistically significant. The fixed effect of Time was statistically significant, $F(1, 213) = 10.62$, $p = .001$. The estimated marginal mean rate of Weaken-Other evidence use increased from pre-test ($EMM = 0.09$, $SE = 0.02$) to post-test ($EMM = 0.18$, $SE = 0.02$), indicating that, relative to total idea units, both groups comparably increased the proportion devoted to weakening opposing positions following the intervention.

4.7.4.1. Support-other. Students' ability to support opposing perspectives was analyzed using a generalized linear mixed model with a Poisson distribution and log link function, controlling for Class. Neither group differences ($B = 0.13$, $SE = 0.34$, $p = .710$) nor the Group $\times$ Time interaction, $F(1, 213) = 0.16$, $p = .690$, were statistically significant. A significant main effect of Time was found, $F(1, 213) = 9.11$, $p = .003$, with estimated marginal means increasing from pre ($EMM = 0.05$, $SE = 0.01$) to post ($EMM = 0.11$, $SE = 0.02$), confirmed by pairwise contrasts (contrast = 0.06, $SE = 0.02$, $t(213) = 3.40$, $p = .001$). These results suggest that students in both conditions similarly improved in supporting alternative perspectives over time, with no differential effect between groups.

4.7.5. Weaken own (weak_own)

Similar to the intervention topic, evidence used to weaken one's own position exhibited extremely low base rates across conditions, with values at or near zero at both time points. As a result, the Poisson GLMM encountered estimation difficulties, rendering the Time and Class effects non-estimable (0 degrees of freedom).

4.7.6. Source of evidence

4.7.6.1. QA evidence. Students' use of evidence from the question-and-answer activities was analyzed using a generalized linear mixed model with a Poisson distribution and log link function, controlling for Class and essay length (IdeaUnits). Of 233 cases, 224 (96.1%) were included. Using a Bonferroni-corrected significance threshold ($\alpha = 0.025$), the Group $\times$ Time interaction was not statistically significant, $F(1, 214) = 0.18$, $p = .675$. A significant main effect of Time was found, $F(1, 214) = 10.26$, $p = .002$, with estimated marginal means increasing from pre ($EMM = 1.40$, $SE = 0.09$) to post ($EMM = 1.93$, $SE = 0.14$), confirmed by pairwise contrasts (contrast = 0.53, $SE = 0.17$, $t(214) = 3.12$, $p = .002$). These results indicate that students in both conditions similarly increased their use of QA-based evidence over time, with no differential effect between groups.

4.7.6.2. Personal evidence. A parallel model examined students' use of evidence drawn from personal knowledge, controlling for Class and essay length (IdeaUnits). Using the Bonferroni-corrected threshold ($\alpha = 0.025$), neither the main effect of Time, $F(1, 214) = 0.58$, $p = .447$, nor the Group $\times$ Time interaction, $F(1, 214) = 1.70$, $p = .194$, reached significance. Estimated marginal means for Time were similar at pre ($EMM = 1.18$, $SE = 0.10$) and post ($EMM = 1.06$, $SE = 0.11$). These results suggest that students' reliance on personal evidence remained stable across time in both conditions.

4.7.7. However arguments

Students' use of however clauses was analyzed using a generalized linear mixed model with a Poisson distribution and log link function, controlling for Class. The Group $\times$ Time interaction was not statistically

significant, $F(1, 213) = 1.05, p = .307$. A significant main effect of Time was found, $F(1, 213) = 10.19, p = .002$, with estimated marginal means increasing from pre (EMM = 0.01, $SE = 0.01$) to post (EMM = 0.07, $SE = 0.01$), confirmed by pairwise contrasts (contrast = 0.06, $SE = 0.01, t(213) = 4.01, p < .001$). These results indicate that students in both conditions similarly increased their use of however clauses over time, with no differential effect between groups.

4.7.8. Reference to both positions

Descriptively, the proportion of students who referenced both positions increased from 19.6% at pre-test to 58.0% at post-test. A generalized linear mixed model (GLMM) with a binomial distribution and logit link function was significant, $F(9, 214) = 5.25, p < .001$, with a classification accuracy of 75.4%. Pairwise comparisons of the estimated marginal means confirmed a significant change from pre-test to post-test (contrast estimate = 0.17, $t = 1.99, p = .048$, 95% CI [0.00, 0.33]). The Time $\times$ Group interaction was not significant, $F(1, 214) = 0.04, p = .841$, indicating that the increase did not differ between the US and Soda groups.

4.8. Qualitative analysis of arguments on the transfer topic

Further to the function-based analysis, of interest is a content-based analysis of the arguments that students developed. Table 4 shows the most frequently used pieces of knowledge at each assessment point along with the frequency of usage and whether that piece of knowledge was included in the Q&A cards or not (marked as given), while the full table is in the Appendix. The majority of the pieces of knowledge used were not included in the Q&A cards, probably arose from students' personal knowledge or as a result of their engagement in discussion, in the case of the final assessment.

The fact that at both the initial and final assessment the most frequently used pieces of knowledge were those conveying Greek Cypriots' sense of belonging to and possession of the houses and land that Turkish settlers now inhabit, as well as those expressing emotions of injustice and highlighting the illegal presence of Turkish settlers (see Table 4) demonstrates that today's elementary school students still hold strong beliefs about the Turkish settlers and the occupation, offering further evidence of the ecological validity of our study.

Qualitative analysis complements quantitative analysis, showing that students used two times more different knowledge claims at post assessment compared to initial assessment. Students used three times more knowledge claims supporting the position of Turkish settlers to leave the houses they are currently living in at post-test ($n = 27$) compared to pre-test ($n = 12$), and two times more knowledge claims supporting the position to stay in the houses at post-test ($n = 17$) compared to pre-test ($n = 9$). The usage of some knowledge claims, which were not provided to them, at the final-assessment such as "It would be unfair, wrong, cruel to make them (i.e., the Turkish Settlers) leave now" and "It would be difficult for them (i.e., the Turkish Settlers)

Table 4

Most Frequently Used Knowledge Claims on the Topic of Turkish Settlers.

Knowledge claims in favor of Turkish Settlers <i>moving out</i> of the houses			
Initial Assessment		Final Assessment	
Idea of belonging & possession- The houses and the land belong to (the) Greek Cypriot(s) (refugees). (given)	72	Idea of belonging & possession- The houses and the land belong to (the) Greek Cypriot(s) (refugees). (given)	90
It is wrong, unjust, unfair and illegal to stay there and not allow the CG to return- they must leave. They are illegal. (only the part that they should return the houses to the legal owners was given)	59	It is wrong, unjust, unfair and illegal to stay there and not allow the CG to return- they must leave. They are illegal. (only the part that they should return the houses to the legal owners was given)	78
Greek Cypriots paid for those houses/they worked (hard) to build them—not the Turkish settlers.(not given)	24	Greek Cypriots have the right to return/They want to return/ They can't forget those houses and their lives there. (given)	28
The way these houses were obtained was illegal and/or violent. (not given)	19	The way these houses were obtained was illegal and/or violent. (not given)	21
They should leave because the houses are old, and they can do better. (not given)	6	Greek Cypriots paid for those houses/they worked (hard) to build them—not the Turkish settlers.	15

Knowledge claims in favor of Turkish Settlers <i>staying</i> in the houses			
Initial Assessment		Final Assessment	
The right of the Turkish Settlers to stay because they have stayed there for so long, have had children and built their life there. (given)	28	The right to stay because the Turkish Settlers have stayed there for so long, have had children and built their life there. (given)	56
It would be unfair, wrong, cruel to make them (i.e., the Turkish Settlers) leave now. (not given)	9	It would be unfair for them (Turkish Settlers) to leave because that would come in opposition to what Turkey promised them in order to leave their home country back in 1974. (given)	20
Greek Cypriots have better houses now—they wouldn't want to move. (not given)	3	It would be unfair, wrong, cruel to make them (i.e., the Turkish Settlers) leave now.	17
They should stay because they came for a better life here. (given)	3	It would be difficult for them (i.e., the Turkish Settlers) to leave now. (not given)	11
For the Turkish Cypriot children, Cyprus is their homeland. (not given)	1	The Greek Cypriots have better houses now—they wouldn't want to leave them. (not given)	4

to leave now" shows that students could see the other side's point of view, providing further support to the findings of the quantitative analysis presented above suggesting that participants improved in their ability to engage in two-sided reasoning on the transfer topic.

Table 3

Means, Standard Deviations, and 95% Confidence Intervals for Idea Units, Evidence, and Types of Evidence Function on the Transfer Topic, by Condition and Time.

	Far-Transfer Condition (n = 53)				Near-Transfer Condition (n = 59)			
	Pre-Test		Post-Test		Pre-Test		Post-Test	
IdeaUnits	M (SD)	95% CI	M (SD)	95% CI	M (SD)	95% CI	M (SD)	95% CI
Evidence	2.30 (1.28)	[1.95, 2.66]	5.04 (2.23)	[4.39, 5.70]	2.77 (1.82)	[2.28, 2.26]	5.11 (2.61)	[4.41, 5.81]
Functional_Ev	2.00 (1.04)	[1.69, 2.31]	4.32 (2.14)	[3.69, 4.95]	2.39 (1.72)	[1.93, 2.85]	4.48 (2.35)	[3.85, 5.11]
Supp_Own	2.00 (1.04)	[1.69, 2.31]	4.13 (2.19)	[3.48, 4.77]	2.34 (1.76)	[1.87, 2.81]	4.29 (2.43)	[3.64, 4.94]
Weaken_Other	1.66 (1.01)	[1.36, 1.96]	2.38 (1.36)	[1.98, 2.78]	1.79 (1.44)	[1.40, 2.17]	2.39 (1.45)	[2.61, 2.78]
Sup_Other	0.02 (0.50)	[0.05, 0.34]	1.19 (1.45)	[0.77, 1.62]	0.38 (0.91)	[0.13, 0.62]	1.21 (1.59)	[0.79, 1.64]
Weaken_Own	0.13 (0.34)	[0.03, 0.23]	0.60 (0.77)	[0.37, 0.82]	0.18 (0.64)	[0.01, 0.35]	0.70 (1.03)	[0.42, 0.97]
Personal	0.00 (0.00)	[0.00, 0.00]	0.00 (0.00)	[0.00, 0.00]	0.00 (0.00)	[0.00, 0.00]	0.00 (0.00)	[0.00, 0.00]
QA_Ev	0.94 (0.92)	[0.67, 1.21]	1.62 (1.36)	[1.22, 2.02]	1.09 (1.21)	[0.77, 1.41]	1.57 (1.45)	[1.18, 1.06]
However	1.06 (0.90)	[0.80, 1.33]	2.53 (1.72)	[2.03, 3.04]	1.25 (0.94)	[1.00, 1.50]	2.73 (1.94)	[2.21, 3.25]
	0.02 (0.15)	[-0.02, 0.64]	0.36 (0.57)	[0.20, 0.53]	0.07 (0.26)	[0.00, 0.14]	0.41 (0.68)	[0.23, 0.59]

5. Discussion

The present study demonstrates the power of engaging in dialogic argumentation to support the development of reasoning skills about real-life, controversial societal issues with personal stakes, a fundamental skill for democratic citizenship. We examined whether argumentation skills developed through an intervention based on dialogic argumentation and reflection, on a neutral topic, transfer to a controversial societal problem of personal relevance. We hypothesized that neutral topics, which do not carry the burden of heightened emotional reactivity encountered with emotionally laden topics that impairs complex cognitive processing (Mason et al., 2017; Mason et al., 2019), could provide favorable conditions for developing reasoning skills. We further hypothesized that once developed, these reasoning skills could transfer to more emotionally charged controversial issues. To test this hypothesis, we followed a sample of young adolescents who over a period of weeks engaged deeply with a neutral topic, either with the topic of US immigrants or the topic of the consumption of soda. Activities in the two conditions were identical except for the topic discussed.

In line with prior empirical work involving this intervention program, both groups exhibited argumentation skill gains on the intervention topic (see Iordanou & Rapanta, 2021, for review). In particular, students in both conditions constructed richer arguments—with more idea units—and more evidence. Students incorporated evidence from the Q&A cards that were made available to them, confirming that providing information in the form of Q&A is an effective method for acquiring new pieces of evidence and incorporating them into one's arguments (Iordanou & Kuhn, 2025). Notably, students employed evidence not only for supporting their own position but most importantly for weakening the opposing position, a more advanced and challenging argument skill, which suggests that students shifted attention from their own to the other's position and were trying to address the other position's arguments. Further supporting this movement from own to other position is the finding showing a statistically significant increase in the proportion of students who referenced both positions from initial to final position. One may claim that this finding may be the result of being exposed to the arguments of the other position, as they were expressed by their peers with whom they had dialogic argumentation, who genuinely represent and support the other position, a condition that previous research shows is effective for developing argumentation skills on a particular topic (Iordanou & Kuhn, 2020). The true value of the intervention, however, would be evidenced by the transfer of gains to a novel context on which they had not received extensive intervention, to rule out the possibility that counterarguments resulted from merely internalizing the opposing arguments to which they were exposed. We turn next to the question of transfer, namely: Do argumentation skills developed through an intervention based on dialogic argumentation and reflection, on a neutral topic, transfer to a controversial real-life problem of personal relevance?

Results regarding the effectiveness of the intervention in supporting the development of students' argumentation skills which are transferable to real-world complex issues offer evidence of both near and far transfer from the intervention topic to a real-life controversial topic of personal relevance. Examining performance on the criteria that we set, which lie at the heart of advanced argumentation skills and are fundamental for democratic citizenship—namely, evidence use and acknowledgment of alternative positions—our findings demonstrate that gains did transfer from the intervention to the transfer topic in both experimental conditions, near- and far-transfer conditions. Students in both conditions showed statistically significant gains in the usage of counterarguments from initial to final assessment on the transfer topic, accompanied by usage of more relevant evidence that was provided to them through Q&A cards. Furthermore, a significant increase in the proportion of students who referenced both positions was observed, moving from a small minority—20%—referencing both positions at initial assessments to the majority—almost 60%—doing so at the final

assessment. Complementary qualitative analysis provided further support for the quantitative findings of gains in constructing evidence-based arguments and engaging in two-sided reasoning on the transfer topic.

Skill gains in argumentation on a new topic are significant because they suggest that something was gained that went beyond the specific topic studied. Skilled arguers have been shown to appreciate the purpose of argumentation and evidence more than novices do (Felton & Crowell, 2022). In the present study, once they gained this appreciation in the context of a particular topic, they were able to also see it in the context of a new topic, and, interestingly, even in a controversial topic on which they held strong emotions and beliefs in favor of one position (Iordanou et al., 2020). Although transfer has been demonstrated in the literature across topics, domains, and media of communication (see Iordanou & Rapanta, 2021, for a review), the novel contribution of the present study is in showing transfer to a personal topic on which participants hold strong beliefs and emotions. One may claim that students could transfer some of the arguments they developed for the US immigrant topic to the transfer topic. Although this alternative explanation cannot be ruled out for the near-transfer condition (US immigration), it fails to account for the far-transfer gains observed in the soda consumption condition. The latter demonstrated remarkable transfer from a neutral science topic to an emotionally and belief-laden sociopolitical issue with minimal topical overlap.

What then accounts for the transfer of gains, especially among the far-transfer condition? The answer, in our view, lies in the development of a meta-level understanding and appreciation of the importance of taking the other's position into account and using evidence to support one's claims when arguing, an understanding that, once developed, appears to be transferable from dialogic argumentation to individual argumentation when constructing an essay and across different contexts. This meta-level understanding, previous work using the microgenetic method has shown, develops gradually alongside the development of argumentation strategies (see Iordanou, 2022; Kuhn et al., 2008). Being prompted by their interlocutors to respond to their arguments under the commonly observed comments of “You haven't answered my question” and through the reflection sheets that explicitly asked them to reflect on whether they had addressed the other position's arguments, at least implicitly conveyed to the experimental participants the suggestion that paying attention to the other's position is important. These discourse norms—such as paying attention to and addressing other's position—which develop gradually during extensive practice in argumentation (Kuhn & Zillmer, 2015; Shi, 2020) are then internalized and implemented in different contexts, such as when writing about an issue after reading multiple sources on it (Iordanou & Fotiou, 2025) and, as the present study shows, when addressing in writing a controversial social issue.

The provision of Q&A cards on the transfer topic in session 9 may have contributed to the observed gains by increasing students' familiarity with the topic and providing them with additional evidence to incorporate into their arguments. Previous research has shown that exposure to the narrative of the “other” side, as authentically expressed by members of that side, contributes to acquiring a multiperspectival understanding of the topic (Iordanou et al., 2025). While we cannot fully disentangle the contribution of increased topic knowledge from the development of transferable argumentation skills, the gains observed in counterargument use and acknowledgment of both positions suggest that the gains reflect something beyond mere information acquisition.

5.1. Limitations and future research

While our findings demonstrate that engagement in argumentation with a neutral topic can serve as a productive context for developing argumentation skills that transfer to a societal controversial issue of personal relevance, the current research design does not allow us to infer whether gains on the controversial topic are attributable to the neutrality of the intervention topics or to engagement in the

argumentation intervention more broadly. Future research should include a comparison condition involving argumentation practice on controversial topics to isolate the specific contribution of topic neutrality. Furthermore, the multi-component intervention implemented in the current study does not allow us to isolate the unique contribution of its individual elements—such as engagement in argumentation with an alternative viewpoint, participation in reflective activities, and acquisition of greater knowledge about the opposing position. Future research could examine and compare the distinct contribution of each component, as well as their combined effect on the observed outcomes, through experimental design. Additionally, including a control group that does not have the opportunity to argue on the transfer topic would help more clearly demonstrate whether the observed gains are attributable to the development of argumentation skills or to increased familiarity with the topic.

It would also be valuable to explore the effectiveness of such an intervention with older populations, who may hold more entrenched my-side beliefs, be more emotionally invested in these issues, and perceive them as integral to their identities. Finally, future studies could examine how direct exposure to opposing arguments—through authentic dialogue with individuals who genuinely hold the alternative position (Iordanou & Kuhn, 2020)—affects students' argumentation skills and beliefs. A particularly promising avenue would be to investigate whether argumentation with AI assuming the opposing position, which offers a safe environment for engagement with the opposing view (Iordanou & Chinn, 2026), could prove equally effective in overcoming the physical and socio-political barriers inherent in facilitating authentic cross-community discussions.

5.2. Educational implications

Our work makes several contributions to the literature. Firstly, our work extends research on the development of argumentation skills by demonstrating transfer from a neutral topic to a social, current, controversial, and affectively laden topic of personal relevance, suggesting that the development of appreciation—and subsequent adherence—of argumentative norms is transferable. Secondly, our work has broad educational implications, adding to the growing body of scholarship that links argumentation to democratic education (e.g., Joshi, 2016; Kuhn, 2025; Moshman, 2020; Rapanta & Iordanou, 2023; Weinstock, 2005), highlighting its potential not only as a critical thinking skill but as a mode of civic formation. Our unique contribution here is suggesting that starting from a neutral topic to develop students' argumentation skills seems a promising way to develop the argument skills needed for controversial social and sociopolitical issues. At the same time, we recognize that classroom engagement with controversial issues inevitably differs from participation in public civic life (Hess, 2009). Educational settings provide pedagogically structured spaces in which complexity can be introduced progressively, evidence can be scaffolded, and dialogue can be facilitated in ways that are seldom possible in everyday political encounters. Rather than viewing this as a limitation alone, we regard such pedagogical structuring as an essential condition for cultivating the argumentative norms and metacognitive dispositions that may later support participation in less structured, emotionally charged, and socially consequential public debates. Whether and under what conditions these competencies transfer to civic engagement beyond the classroom remains an important question for future research. Finally, our work carries important implications for the teaching of controversial issues in conflict-affected societies such as Cyprus. In particular, the findings emphasize the significance of integrating argumentation skills within peace pedagogies that engage controversial issues (Bekerman & Zembylas, 2012; King, 2009; Zembylas & Kambani, 2012). The study points to the need for pedagogical approaches that foster argumentation skills in ways that are sensitive to students' emotional investments in social, sociopolitical, and historical issues, while simultaneously nurturing values and dispositions

to peace.

In conclusion, the present work highlights that engaging students in dialogic argumentation helps them resist simplistic narratives, which often result from my-side biases, by encouraging deeper engagement with evidence, complexity, and multiperspectivity (Iordanou et al., 2025); these skills are the cornerstone of democratic citizenship. While schools cannot fully reproduce the moral urgency, emotional intensity, and political complexity of public controversies, they can provide protected spaces in which learners develop the epistemic and dialogic capacities needed to engage such controversies more thoughtfully. In this sense, our findings identify a crucial educational pathway through which schools can begin to cultivate the forms of reasoning that democratic participation requires.

CRedit authorship contribution statement

Kalypso Iordanou: Writing – original draft, Supervision, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Constantina Fotiou:** Writing – review & editing, Investigation, Formal analysis. **Michalinos Zembylas:** Writing – review & editing, Supervision, Funding acquisition, Conceptualization.

Declaration of competing interest

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

Acknowledgment

This research is funded by the European Commission and the Republic of Cyprus, through a grant received by the Research & Innovation Foundation, Excellence hubs No: EXCELLENCE/0421/0500.

AI declaration

During the preparation of this work, the authors used Claude in order to proofread segments of the manuscript. After using this tool/service, the author(s) reviewed and edited the content as needed and take full responsibility for the content of the published article.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.cedpsych.2026.102486>.

Data availability

Data will be made available on request.

References

- Bail, C. A., Argyle, L. P., Brown, T. W., Bumpus, J. P., Chen, H., Hunzaker, M. B. F., ... Volfovsky, A. (2018). Exposure to opposing views on social media can increase political polarization. *Proceedings of the National Academy of Sciences*, 115(37), 9216–9221.
- Barberá, P., Jost, J. T., Nagler, J., Tucker, J. A., & Bonneau, R. (2015). Tweeting from left to right: Is online political communication more than an echo chamber? *Psychological Science*, 26(10), 1531–1542. <https://doi.org/10.1177/09567976155594>
- Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn?: A taxonomy for far transfer. *Psychological Bulletin*, 128(4), 612–637. <https://doi.org/10.1037/0033-2909.128.4.612>
- Barzilai, S., & Weinstock, M. (2020). Beyond trustworthiness: Comprehending multiple source perspectives. In P. Van Meter, A. List, D. Lombardi, & P. Kendeou (Eds.), *Handbook of learning from multiple representations and perspectives* (pp. 123–140). Routledge/Taylor & Francis Group. <https://doi.org/10.4324/9780429443961-11>.
- Bekerman, Z., & Zembylas, M. (2012). *Teaching contested narratives: Identity, memory and reconciliation in peace education and beyond*. Cambridge University Press.

- Binnendyk, J., & Pennycook, G. (2022). Intuition, reason, and conspiracy beliefs. *Current Opinion in Psychology*, 47, Article 101387. <https://doi.org/10.1016/j.copsyc.2022.101387>
- Christodoulou, E., & Iordanou, K. (2021). Democracy under attack: Challenges of addressing ethical issues of AI and big data for more democratic digital media and societies. *Frontiers in Political Science*, 3, Article 682945. <https://doi.org/10.3389/fpos.2021.682945>
- Christou, M. (2007). The language of patriotism: Sacred history and dangerous memories. *British Journal of Sociology of Education*, 28(6), 709–722. <https://doi.org/10.1080/01425690701609946>
- Felton, M., & Crowell, A. (2022). Argumentation as a Collaborative Enterprise: A study of dialogic purpose and dialectical relevance in novice and experienced arguers. *Informal Logic*, 42(1), 171–202.
- Firetto, C. M., Murphy, P. K., Starrett, E., Herman, E. A., Greene, J. A., Tang, Y., & Yan, L. (2025). Investigating grade-level and text genre effects in quality talk discussions: An AI-powered discourse analysis of upper primary students' high-level comprehension. *Learning and Instruction*, 100, Article 102208. <https://doi.org/10.1016/j.learninstruc.2025.102208>
- Georgiou, N., Delfabbro, P., & Balzan, R. (2021). Conspiracy theory beliefs, scientific reasoning and the analytical thinking paradox. *Applied Cognitive Psychology*, 35(6), 1523–1534. <https://doi.org/10.1002/acp.3885>
- Hess, D. E. (2002). Discussing controversial public issues in secondary social studies classrooms: Learning from skilled teachers. *Theory & Research in Social Education*, 30(1), 10–41. <https://doi.org/10.1080/00933104.2002.10473177>
- Hess, D. E. (2004). Controversies about controversial issues in democratic education. *PS: Political Science & Politics*, 37(2), 257–261. <https://doi.org/10.1017/S1049096504004196>
- Hess, D. E. (2009). *Controversy in the classroom: The democratic power of discussion*. Routledge.
- Institute for Economics & Peace. (2025). Global peace index 2025: The great fragmentation. <https://www.economicsandpeace.org/wp-content/uploads/2025/06/Global-Peace-Index-2025-Press-Release.pdf>
- Iordanou, K. (2022). Supporting strategic and meta-strategic development of argument skill: The role of reflection. *Metacognition and Learning*, 17, 399–425. <https://doi.org/10.1007/s11409-021-09289-1>
- Iordanou, K., & Chinn, C. A. (2026). *Arguing with GenAI: A framework for productive human-GenAI interaction* [Manuscript submitted for publication].
- Iordanou, K., & Constantinou, C. P. (2015). Supporting use of evidence in argumentation through practice in argumentation and reflection in the context of SOCRATES learning environment. *Science Education*, 99, 282–311. <https://doi.org/10.1002/sce.21152>
- Iordanou, K., & Fotiou, C. (2025). Supporting integration of multiple source perspectives through dialogic argumentation. *Frontline Learning Research*, 13(1), 22–44. <https://doi.org/10.14786/flr.v13i1.1241>
- Iordanou, K., Fotiou, C., Manoli, A., & Zembylas, M. (2025). Examining myside bias on a controversial historical event after engagement in dialogic argumentation: Insights from a think aloud study. *Learning and Instruction*, 100, Article 102209. <https://doi.org/10.1016/j.learninstruc.2025.102209>
- Iordanou, K., Kendeou, P., & Zembylas, M. (2020). Examining my-side bias during and after reading controversial historical accounts. *Metacognition and Learning*, 15(3), 319–342. <https://doi.org/10.1007/s11409-020-09240-w>
- Iordanou, K., & Kuhn, D. (2020). Contemplating the opposition: Does a personal touch matter? *Discourse Processes*, 57(4), 343–359. <https://doi.org/10.1080/0163853X.2019.1701918>
- Iordanou, K., & Kuhn, D. (2025). Investigating climate change through argumentation: Purposeful questioning supports argumentation and knowledge acquisition. *Journal of Experimental Psychology: Applied*. <https://doi.org/10.1037/xap0000534>. Advance online publication.
- Iordanou, K., Kuhn, D., Matos, F., Shi, Y., & Hemberger, L. (2019). Learning by arguing. *Learning and Instruction*, 63, Article 101207. <https://doi.org/10.1016/j.learninstruc.2019.05.004>
- Iordanou, K., & Rapanta, C. (2021). “Argue with me”: A method for developing argument skills. *Frontiers in Psychology*, 12, Article 631203. <https://doi.org/10.3389/fpsyg.2021.631203>
- Joshi, P. (2016). Argumentation in democratic education: The crucial role of values. *Theory Into Practice*, 55(4), 279–286. <https://doi.org/10.1080/00405841.2016.1208066>
- Kahan, D. M. (2013). Ideology, motivated reasoning, and cognitive reflection. *Judgment and Decision making*, 8(4), 407–424. <https://doi.org/10.1017/S1930297500005271>
- Kerr, D., & Huddelstone, T. (2015). *Living with controversy: Teaching controversial issues through education for democratic citizenship and human rights (EDC/HRE)*. Council of Europe. <https://rm.coe.int/16806948b6>
- King, J. T. (2009). Teaching and learning about controversial issues: Lessons from Northern Ireland. *Theory & Research in Social Education*, 37(2), 215–246. <https://doi.org/10.1080/00933104.2009.10473395>
- Klaczynski, P. A. (2000). Motivated scientific reasoning biases, epistemological beliefs, and theory polarization: A two-process approach to adolescent cognition. *Child Development*, 71(5), 1347–1366. <https://doi.org/10.1111/1467-8624.00232>
- Korteling, J. E., Brouwer, A.-M., & Toet, A. (2018). A neural network framework for cognitive bias. *Frontiers in Psychology*, 9, 1561. <https://doi.org/10.3389/fpsyg.2018.01561>
- Kubin, E., & von Sikorski, C. (2021). The role of (social) media in political polarization: A systematic review. *Annals of the International Communication Association*, 45(3), 188–206. <https://doi.org/10.1080/23808985.2021.1976070>
- Kuhn, D. (2025). What can discourse do? *Topoi*, 1–11. <https://doi.org/10.1007/s11245-025-10224-1>
- Kuhn, D., Goh, W., Iordanou, K., & Shaenfield, D. (2008). Arguing on the computer: A microgenetic study of developing argument skills in a computer-supported environment. *Child Development*, 79(5), 1310–1328. <https://doi.org/10.1111/j.1467-8624.2008.01190.x>
- Kuhn, D., Hemberger, L., & Khait, V. (2016). Tracing the development of argumentative writing in a discourse-rich context. *Written Communication*, 33(1), 92–121. <https://doi.org/10.1177/0741088315617157>
- Kuhn, D., & Iordanou, K. (2022). Why do people argue past one another rather than with one another? In D. Dunning, & N. Ballantyne (Eds.), *Reason, bias, and inquiry: New perspectives from the crossroads of epistemology and psychology*. Oxford University Press. <https://doi.org/10.1093/oso/9780197636916.003.0015>
- Kuhn, D., & Modrek, A. (2021). Mere exposure to dialogic framing enriches argumentative thinking. *Applied Cognitive Psychology*, 35(5), 1349–1355. <https://doi.org/10.1002/acp.3862>
- Kuhn, D., & Zillmer, N. (2015). Developing norms of discourse. In L. B. Resnick, C. S. C. Asterhan, & S. N. Clarke (Eds.), *Socializing intelligence through academic talk and dialogue* (pp. 77–86). American Educational Research Association.
- Liger, Q., Gutheil, M., Faion, M., Dzekova, R., Stoyanova, N., Savev, M., Ralchev, S., & Comunale, T. (2022). *Right-wing extremism in the EU* (study no. PE 700.953). European Parliament. [https://www.europarl.europa.eu/RegData/etudes/STUD/2021/700953/IPOL_STU\(2021\)700953_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2021/700953/IPOL_STU(2021)700953_EN.pdf)
- Mason, L., Scrimin, S., Tornatora, M. C., & Zaccoletti, S. (2017). Emotional reactivity and comprehension of multiple online texts. *Learning and Individual Differences*, 58, 10–21. <https://doi.org/10.1016/j.lindif.2017.07.002>
- Mason, L., Zaccoletti, S., Scrimin, S., Tornatora, M. C., Florit, E., & Goetz, T. (2019). Reading with the eyes and under the skin: Comprehending conflicting digital texts. *Journal of Computer Assisted Learning*, 36(2), 162–174. <https://doi.org/10.1111/jcal.12399>
- Moshman, D. (2020). *Reasoning, argumentation, and deliberative democracy* (1st ed.). Routledge. <https://doi.org/10.4324/9780429316029>
- Moudouros, N. (2025). Settler colonialism or a hybrid case? Dimensions of colonization in Cyprus and Turkish Cypriot-settler antagonism. *International Journal of Anthropology and Ethnology*, 9(14). <https://doi.org/10.1186/s41257-025-00137-7>
- Mudde, C. (2019). *The far right today*. Polity Press.
- Murphy, P. K., & Firetto, C. M. (2018). Quality talk: A blueprint for productive talk. In P. K. Murphy (Ed.), *Classroom discussions in education: Promoting productive talk about text and content* (pp. 101–133). Routledge.
- Papadakis, Y. (2008). Narrative, memory and history education in divided Cyprus: A comparison of schoolbooks on the “history of Cyprus”. *History and Memory*, 20(2), 128–148. <https://doi.org/10.2979/his.2008.20.2.128>
- Pew Research Center. (2023). Americans’ feelings about politics, polarization, and the tone of political discourse. , September 19 <https://www.pewresearch.org/politics/2023/09/19/americans-feelings-about-politics-polarization-and-the-tone-of-political-discourse/>.
- Piaget, J. (1932). *The moral judgment of the child*. Kegan Paul, Trench, Trubner & Co.
- Rapanta, C., & Felton, M. K. (2022). Learning to argue through dialogue: A review of instructional approaches. *Educational Psychology Review*, 34, 477–509. <https://doi.org/10.1007/s10648-021-09637-2>
- Rapanta, C., & Iordanou, K. (2023). Argumentation and critical thinking. In R. J. Tierney, F. Rizvi, & L. Ercikan (Eds.), *International encyclopedia of education* (4th ed., pp. 575–587). Elsevier. <https://doi.org/10.1016/B978-0-12-818630-5.14073-4>
- Shi, Y. (2019). Enhancing evidence-based argumentation in a mainland China middle school. *Contemporary Educational Psychology*, 59, Article 101809. <https://doi.org/10.1016/j.cedpsych.2019.101809>
- Shi, Y. (2020). Talk about evidence during argumentation. *Discourse Processes*, 57(9), 770–792. <https://doi.org/10.1080/0163853X.2020.1777498>
- Spyrou, S. (2006). Constructing “the Turk” as an enemy: The complexity of stereotypes in children’s everyday worlds. *South European Society & Politics*, 11(1), 95–110. <https://doi.org/10.1080/13608740500470364>
- Stanovich, K. E., & West, R. F. (2007). Natural myside bias is independent of cognitive ability. *Thinking & Reasoning*, 13(3), 225–247. <https://doi.org/10.1080/13546780600780796>
- Stanovich, K. E., West, R. F., & Toplak, M. E. (2013). Myside bias, rational thinking, and intelligence. *Current Directions in Psychological Science*, 22(4), 259–264. <https://doi.org/10.1177/0963721413480174>
- U.S. Census Bureau. (2017). High school completion rate is highest in U.S. history [press release]. <https://www.census.gov/newsroom/press-releases/2017/educational-attainment-2017.html>
- Weinstock, M. P. (2005). Cognitive bases for effective participation in democratic institutions: Argument skill and juror reasoning. *Theory & Research in Social Education*, 33(1), 73–102. <https://doi.org/10.1080/00933104.2005.10473272>
- Zembylas, M. (2013). The emotional complexities of “our” and “their” loss: The vicissitudes of teaching about/for empathy in a conflicting society. *Anthropology & Education Quarterly*, 44(1), 19–37. <https://doi.org/10.1111/J.1548-1492.2012.01175>
- Zembylas, M. (2021). *Affect and the rise of right-wing populism: Pedagogies for the renewal of democratic education*. Cambridge University Press.
- Zembylas, M. (forthcoming 2026). Liberal epistemologies of peace and the problem of innocence: A decolonial critique of peace education. In Hussak, M., & Harbeck-Pingel, B. (Eds.), *Global perspectives on peace education*. Barbara Budrich Publishers.
- Zembylas, M., Charalambous, C., & Charalambous, P. (2016). *Peace education in a conflict-troubled society: An ethnographic journey*. Cambridge University Press.
- Zembylas, M., & Kambani, F. (2012). The teaching of controversial issues during elementary-level history instruction: Greek-Cypriot teachers’ perceptions and

- emotions. *Theory & Research in Social Education*, 40(2), 107–133. <https://doi.org/10.1080/00933104.2012.670591>
- Zembylas, M., & Loukaidis, L. (2021). Affective practices, difficult histories and peace education: An analysis of teachers' affective dilemmas in ethnically divided Cyprus. *Teaching and Teacher Education*, 97, 1–11. <https://doi.org/10.1016/j.tate.2020.103225>
- Zembylas, M., Loukaidis, L., & Antoniou, P. (2020). Teachers' understandings of empathy in teaching about the holocaust in Cyprus: The emotional risks of identification and the disruptive potential of "empathic unsettlement.". *Teaching and Teacher Education*, 89, Article 103007. <https://doi.org/10.1016/j.tate.2019.103007>
- Zembylas, M., & Michaelidou, A. (2013). Exploring children's perceptions of forgiveness at the interpersonal and political levels: Emerging evidence in Cyprus. *Education 3-13*, 41(3), 265–281. <https://doi.org/10.1080/03004279.2011.574641>
- Zmigrod, L., Rentfrow, P. J., & Robbins, T. W. (2019). Cognitive inflexibility predicts extremist attitudes. *Frontiers in Psychology*, 10, 989. <https://doi.org/10.3389/fpsyg.2019.00989>