



Research paper

Can teachers implement a student-centered dialogical argumentation method across the curriculum?



Chrysi Rapanta*

Universidade Nova de Lisboa, Portugal

HIGHLIGHTS

- Teachers and not researchers implemented a dialogical argumentation curriculum.
- Their treatment classes did better than their control classes in argumentative writing.
- This improvement was evident in four disciplinary areas.
- Teachers reported additional metacognitive and socio-emotional students' gains.
- Teachers' implementation of the curriculum was easy and effective.

ARTICLE INFO

Article history:

Received 19 January 2021
 Received in revised form
 24 May 2021
 Accepted 28 May 2021
 Available online xxx

Keywords:

Argument-based teaching
 Student-centered learning
 Discourse
 Writing skills
 Dialogical argumentation

ABSTRACT

Does student-centered dialogical argumentation work in authentic classroom settings? Four experienced secondary teachers implemented an argument-based curriculum situated within the contexts of their four respective disciplines - language arts, history, science, and civic education. A mixed-method analysis showed that students who participated in the curriculum performed significantly better on a final argumentative essay, compared to control groups who studied the same content with the same teachers using customary methods. Teachers' interviews suggested additional effects including cognitive, meta-cognitive and socio-emotional skills. Discussion focuses on teachers' role within a dialogical curriculum and argumentation's applicability and efficiency as a transdisciplinary pedagogical method.

© 2021 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Dialogical argumentation has gained terrain in recent years as a learning method, shown to lead to significant improvements in students' oral and written communication, and critical thinking (Chen, Hand, & Park, 2016; Crowell & Kuhn, 2014; Kuhn & Crowell, 2011; Reznitskaya et al., 2001; Venville & Dawson, 2010; Zohar & Nemet, 2002). Extensive research with students of middle and secondary school age has led to the conclusion that the practice of dialogical argumentation helps adolescents to acquire critical thinking skills, most notably argument and counterargument construction, and claim-evidence coordination (Alexander, 2008; Crowell & Kuhn, 2014; Hemberger, Kuhn, Matos, & Shi, 2017; Kuhn

& Crowell, 2011; Kuhn, Hemberger, & Khait, 2016a; Nussbaum & Schraw, 2007; Reznitskaya et al., 2012). Such achievements are major goals of education, especially in the present era of fake news and alternative facts. The key question asked in the work reported here is this: Does dialogical argumentation achieve its intended goals in authentic practice, i.e., when implemented by teachers in ordinary classroom contexts?

The implementation of dialogical argumentation as a pedagogical method, also called argument-based teaching, is different than dialogic teaching (Hardman, 2019; Sedova, Sedlacek, & Svaricek, 2016), although the same principles of democratic talk and knowledge co-construction apply. A main difference lies in whether students' improvement of argumentation skills is an explicit goal of the pedagogies used (that will be the case of argument-based teaching) or a possible by-product of the productive discussions held with peers and the teacher (that will be the case of dialogic teaching). Although there is much research confirming the use and

* Faculty of Social Sciences and Humanities, Universidade Nova de Lisboa, Avenida de Berna 26, 1069-061, Lisbon, Portugal.
 E-mail address: crapanta@fch.unl.pt.

effectiveness of the latter approach in the so-called 'dialogic classrooms' (Christoph & Nystrand, 2001), there is still a need to understand how teachers' implementation of argument-based activities or curricula is possible in the classroom (Larrain et al., 2017). This need is stressed by the fact that argumentation, understood as a socio-cognitive activity in which knowledge and ideas are confronted to each other until a best explanation is reached (Rapanta & Macagno, 2019), is a key competence for 21st century learners.

2. Theoretical background

Despite the evidence of the pedagogical impact and potential of dialogical argumentation (Chen & al., 2016; De La Paz, Ferretti, Wissinger, Yee, & MacArthur, 2012; Duschl & Osborne, 2002; Evagorou, Jimenez-Aleixandre, & Osborne, 2012; Iordanou, Kuhn, Matos, Shi, & Hemberger, 2019; Reznitskaya & Gregory, 2013), teachers are often hesitant to implement it, either because they think it may bring conflict and competition among students (Ferretti & Lewis, 2013), or because they consider argumentation as an "extra-curricular" activity and not as a communication competence transversal to different contexts and, therefore, a necessary part of any curriculum (Wolfe, 2011). Therefore, most of the evidence regarding argument-based teaching has come from researchers rather than by teachers themselves. This is surprising given the increasing attention being given to argumentative essay writing in schools at all levels and the long-standing evidence of uniformly weak expository writing performance (Ferretti & Graham, 2019). Establishing the value of dialogical argumentation as a teacher-implemented pedagogical method in middle and secondary classrooms across the curriculum thus remains a largely unfulfilled goal.

Within a sociocultural approach to learning, classrooms are conceived as sites where the social construction of meanings takes place (Vygotsky, 1978; Wertsch & Toma, 2012). Within a classroom-based activity setting, argumentation is a social practice, mediated by semiotic tools, which leads to "high literacy" outcomes (Newell, Beach, Smith, & VanDerHeide, 2011), manifested in, but not limited to, so-called argumentation skills. These include constructing and identifying valid arguments, supporting arguments with evidence, considering alternative arguments and/or counterarguments, and anticipating and/or replying to counterarguments via rebuttal (Anderson et al., 2001; Rapanta, 2019; Kuhn, 1991).

The "Argue with Me" curriculum (Kuhn, Hemberger, & Khait, 2016b) is a prominent example of a dialogical approach to developing all the argument skills mentioned above. Although a number of argument-based or argument-oriented curricula exist, they are largely discipline-specific. This means that they contain materials and activities that only make sense for the particular disciplinary area for which they were conceived. Examples of such curricula are: the Argument-Driven Inquiry (Sampson, Grooms, & Walker, 2011), the Argue-WISE learning environment (Evagorou et al., 2012), or the Galapagos finches' software-based program (Zemba-Saul, Munford, Crawford, Friedrichsen, & Land, 2002), for science; the Collaborative Reasoning curriculum (Anderson, Chinn, Chang, Waggoner, & Yi, 1997) for language and arts; the "Rosa Parks" (Herrenkohl & Cornelius, 2013) and the "Thinking with historical documents" curricula (De La Paz and Felton, 2010) for history. In contrast to discipline-constrained argument-based curricula, the "Argue with Me" (AWM) curriculum (Kuhn et al., 2016b) is adaptable to any disciplinary field across middle and upper secondary education. A key characteristic of AWM is its commitment to the idea of dialogue and argumentation as a method and an objective for learning, rather than to the disciplinary learning outcomes that may co-occur with this achievement. In this sense, the AWM curriculum is itself a systematic method for the development of

students' argumentative reasoning skills, known in the literature as a 'learning to argue' in contrast to 'arguing to learn' objective (Muller-Mirza & Perret-Clermont, 2009; Von Aufschnaiter, Erduran, Osborne, & Simon, 2008). Other interventions exist that aim toward this goal, but they do not consist of a fully developed curriculum and method.

Given this particular nature of AWM, and its well documented impact on students' argumentative reasoning, discourse, and writing gains (Iordanou & Rapanta, 2021; Crowell & Kuhn, 2014; Hemberger et al., 2017; Iordanou et al., 2019; Kuhn & Crowell, 2011; Kuhn & Moore, 2015; Kuhn et al., 2016a), the present study aims to further explore its transdisciplinary nature, as well as its implementation feasibility by teachers. The idea of transdisciplinary curricula lies in the need for the development of cross-curricular skills, i.e., skills that are important for any disciplinary knowledge field, because of their foundational status in application of one's intellect and intellectual development (Kuhn & Moore, 2015; Shank, 2011; Wolfe, 2011). This is so, given that learning to argue implies acquiring a range of competencies related to informal reasoning and critical thinking (Kuhn, 2018a; Means & Voss, 1996), analytical skills (Glassner & Schwarz, 2005), and metastrategic and epistemological thinking (Kuhn, Zillmer, Crowell, & Zavala, 2013), as well as having implications for democratic citizenship skills, values and dispositions (Kuhn, Feliciano, & Kostikina, 2019).

Within a dialogical argumentation teaching framework, the teacher's role has been studied in terms of discursive strategies, such as the type of questions teachers ask (Larrain, Freire, & Howe, 2014; McNeill & Pimentel, 2010), the explicit focus they place on argumentation goals (Berland & Reiser, 2009), or the time they devote to metacognitive reflection or "deep processing" (Chinn & Brewer, 1998; Duschl & Osborne, 2002). Similarly, the challenges and problems detected in *teachers'* implementation of argument-based teaching (in contrast to *researchers'* implementation) are several. They range from the impact of teachers' beliefs regarding argument-based teaching (Erduran, Simon, & Osborne, 2004; McNeill, 2009; Pimentel & McNeill, 2013) to teachers' challenges in using specific educational materials (Brown, 2009; McNeill et al., 2017), involving issues such as the lack of understanding of argumentation and its components (Sampson & Blanchard, 2012). In addition is the issue of teacher confidence, leading to the hesitation to shift from a teacher-centered, authoritative discourse to a student-centered, collaborative inquiry discourse (McNeill, 2009; McNeill & Pimentel, 2010; Reznitskaya & Gregory, 2013; Wilkinson et al., 2017).

However framed, in order to be sustainable, argumentation-based pedagogies require a shift in the way teachers view both learning itself and their own role in the promotion and assessment thereof (Evagorou & Dillon, 2011; Martin & Hand, 2009). As Zohar (2008) maintains, this shift is not easy, as it mostly relates to teachers' underlying theory of instruction, which needs to be linked to a pedagogy of knowledge construction rather than knowledge transmission. As response to these challenges, and in order to guide teachers in creating argument-oriented classrooms, several materials and curricula have been thus far proposed explicitly aiming at the development of argument skills. The majority of those curricula fall within a science inquiry-based learning approach (e.g., McNeill, 2009; Osborne, Erduran, & Simon, 2004; Zemba-Saul et al., 2002). Fewer approaches explicitly include the interpersonal aspects of argumentation, namely students "directly asking and answering each other's questions as well as commenting on previous ideas in the discussion" (McNeill, 2009, p. 263). The dialogical argumentation curriculum by Kuhn et al. (2016b) is one of the few exceptions; however, its value as a pedagogical method in schools needs further investigation across the curriculum and implemented by classroom teachers.

The present study is intended to address this gap. The goal is to combine a study of teachers' learning and implementation of an argument-based curriculum with the study of the effects thereof on students' argumentation skills in different subject matter areas, namely history, civic education, language arts, and science. This objective was formalized in the following leading research question: *Can teachers implement a student-centered dialogical argumentation method across the curriculum?* Teachers' capacity was judged both from a students' and a teachers' performance perspective as evident in the two study's sub-questions, namely:

- What is the impact, if any, of the AWM curriculum on adolescent students, when the curriculum is implemented by their teachers, in their corresponding subject areas?
- What is the impact, if any, of the AWM curriculum implementation on the teachers themselves, when it comes to their perception of its benefits and/or constraints as well as their role within a student-centered learning environment?

3. Method

3.1. Participants and design

The participants were students from six classes of the lower and two classes of the upper secondary grades of three different public schools in Lisbon, Portugal, with a comparable (middle range) socio-economical profile defined by the average income level of students' families. Their total number was 211, distributed as follows: (a) 51 students from two 7th-grade classes (26 and 25 students correspondingly); (b) 111 students from four 8th-grade classes (28, 30, 27, and 26 students correspondingly); and (c) 49 students from two 10th-grade classes (25 and 24 students correspondingly). In Portugal, students start the 7th grade at 12 years old, the 8th grade at 13 years old, and the 10th grade at 15 years old. The average age of the students at the end of the study was 13.8 years old. Their total gender distribution was 107 girls and 104 boys.

The teachers were three female and one male, all of them in their fifties; the average teaching experience was 25 years (ranging from 22 to 27.5 years). Their teaching subjects were: language arts (Teacher 1), civic education (Teacher 2), science (Teacher 3), and history (Teacher 4). All four teachers had previous experience with using dialogue and argumentation in their classrooms, but not in the specific format of a structured argument-based curriculum. Their participation in the study was voluntary, and they were invited because of their previous motivation and experience.

The quasi-experimental study design required each of the four teachers to teach two classes each, for a total of eight classes. Teacher 1 participated with two 10th-grade classes, Teacher 2 participated with two 7th-grade classes, Teacher 3 participated with two 8th-grade classes, and Teacher 4 with another two 8th-grade classes. Four classes (one taught by each teacher) formed the Treatment Group, henceforth TG, and the other four (one by each teacher) formed the Control Group, henceforth CG. The selection of CG and TG classes was randomly done by the teachers. The two classes per teacher were comparable in terms of number of students, gender distribution, and academic performance. Initial equivalence was confirmed by a pre-test writing task, which yielded no statistically significant differences across each TG/CG class pair ($t_{143} = -0.671, p = .504 > 0.01$). This pre-test writing task was a curriculum-related question chosen and administered by each one of the teachers as part of an ordinary assessment process (i.e. subject matter test).

The participating teachers were informed that after their successful completion of a two-day professional development session

(see below), during the course of the study they would be observed in their TG classrooms by a member of the research team. The four TG classes followed a ten-session argument-based curriculum as detailed below, whereas the CG classes did not. Teachers had been asked earlier about their preferences regarding the specific topics they would like to teach in their respective classes. As Table 1 shows, Teachers 1 and 2 opted for the 'refugees' topic (*Should we receive more refugees?*), Teacher 3 opted for the 'space exploration' topic (*Should we invest more in space exploration?*), whereas Teacher 4 suggested a topic on 'slavery' (*Was slavery a right decision in the 16th century?*). The decisions on topics were made by the teachers. Suggestions were offered by the research team but not limited to them (for example, the 'slavery' topic was entirely a teacher's idea).

Each CG class was exposed to the same topic (Table 1) and its specified contents as its paired TG class during a period of time comparable to the ten sessions of the argument-based curriculum followed only by the TG classes. The common instruction given to the teachers regarding their work with the CG class was "to do what you normally do, strictly avoiding students' engagement in peer-to-peer dialogue or argumentation." To facilitate the process, a teacher-practitioner cooperation approach was followed, in the form of a co-learning agreement (Wagner, 1997). During a 2-h session with each one of the four teachers, the research team discussed with them alternative ways and activities to address the task of exposing the CG students to the same course content available to the TG students but without engaging them in constructive peer-to-peer dialogue and argumentation. Teacher 1 opted for presenting the contents as information the CG students could use to construct solutions for refugees to feel more welcome in their host country; Teacher 2 did a memory exercise with students reading the QAs related to the refugees' topic at home, split into parts, and then the same knowledge was reconstructed in class through roleplaying exercises (some students pretended to be refugees and other students 'played' the locals); Teacher 3 opted for the chair interview technique, in which students had to memorize the information given in the QAs, then pretended to be global policy makers answering questions regarding space exploration opportunities and risks; finally, Teacher 4 organized a whole-class quiz game, again focusing on the QAs previously constructed for the slavery topic.

Overall, all teaching methods in the CG classes focused on individual information search and problem solving, while avoiding any constructive dialogue and argumentation as part of the whole-class discussions or small-group activities. The time dedicated to the in-class activities was five class hours, while an additional 4 h was dedicated to preparatory homework, and 1 h to the in-class writing of the essay. In total, all CG classes dedicated ten class hours for the activity, which is exactly the time spent by the TG classes on the AWM curriculum.

To facilitate the implementation of the argument-based curriculum, teachers were all provided an adapted written lesson plan sequence with the detailed activities and estimated timing for each class, as a guide to follow (see Appendix 1). All TG classes were attended by the author herself and at least one other member of the research team, who closely observed and audio-recorded both the teachers and the students in order to confirm that a faithful implementation of the lesson plan took place. Moreover, a fidelity treatment checklist (see Appendix 2) was completed by the four teachers at the end of the project implementation that confirmed that the approach followed for the CG classes did not share the dialogic, argument-based nature of the AWM lesson plans implemented only with the TG classes.

Table 1
Implementation information.

Teacher	Stud. age	Subject	Essay Topic ^a	Duration ^a
Teacher 1	15	Language arts	Should Portugal receive more refugees?	2 weeks
Teacher 2	13–14	Citizenship education	Same as above.	5 weeks
Teacher 3	12–13	Science	Should we invest more in space exploration?	4 weeks
Teacher 4	13–14	History	Was slavery a right decision in the 16th century?	3 weeks

^a Note: Same for both TG and CG classes.

3.2. Procedure

3.2.1. Teachers' professional development

The four teachers voluntarily participated in two 6-h days of training, for a total of 12 training hours. The following topics were covered: (a) the significance of argumentation; (b) the nature of argument skills; (c) rationale and procedures of the dialogue-based AWM curriculum (Kuhn et al., 2016b); (d) possible topics to work with (Kuhn, 2018b); (e) the various pedagogical activities included (same-side small-group work, opposing-side dialogs, use of evidence; reflection on dialogs; culminating whole-class debate, final individual essay); (f) details of the activities for each of the ten sessions; and (g) anticipated outcomes as reflected in students' writing. The methods used in the training sessions were mixed, alternating between presentation of contents by the instructor and subsequent implementation of the activities by the teachers themselves as if they were students. This experiential aspect of teacher training is usually linked to successful, innovative pedagogies of teacher professional development (Lieberman, 1995).

3.2.2. The argument-based curriculum

The argument-based curriculum, followed by the four TG classes, consisted of a ten-session adaptation of the AWM curriculum (Kuhn, 2018b; Kuhn et al., 2016b). It followed the same structure for all teachers and classrooms. Before the first day of implementation, the four teachers asked their TG students to choose the side they favored (Yes, No, or Uncertain) regarding the main topic (Table 1). Each class was split into two same-side teams having an equal number of members; the minority of students who did not take a clear side (i.e., who chose the option "Uncertain") were distributed between the two teams to equate their size. Each team was further split into subgroups of 4–5 students, balanced from the point of view of both gender and academic ability.

The first two sessions (corresponding to one class hour of 45 min) constituted the "Pregame." Students in their small groups worked on the task of generating reasons supporting their side and identifying evidence in support of such arguments. They were provided post-it notes of different colors and an A5 paper to be used for a group poster. They were also offered short Q&As containing factual information related to the topic that they were told might be useful in their argumentation if they wished to use it (for an example of Q&As used in the AWM curriculum see Rapanta & Trovão, 2020). The same Q&As were provided to both teams and were mixed as to which position they most helped to support.

In the next three sessions ("Game"), groups were further split into dyads, or a dyad and a triad, and paired with a dyad or triad from the opposing side. Each dyad or triad on one side was provided a pad on which to begin a written dialogue with a pair on the other team. The matched pairs were instructed to argue as to why their position was the better one. While waiting for the other side to reply, pairs (or triads) worked on a summary sheet (Kuhn et al., 2016b) asking them to reflect on the two positions. At the next session, each dyad/triad engaged a different opposing dyad/triad.

The remaining five sessions ("Endgame") began with two sessions during which students reassembled into their original teams

to prepare for a final debate. Each group prepared final posters that portrayed their arguments (reasons supported by evidence), possible counterarguments and rebuttals. They also prepared posters summarizing the opposing side's arguments (other-side strengths) and the best counterarguments to weaken them (own-side strengths).

The eighth session was devoted to a final full-class verbal debate, during which a succession of students from each side were paired to debate one-on-one. They were allowed a "huddle" to call on teammates for assistance if needed. During the ninth session, a reflection on the debate took place in a whole-class discussion format. After this reflection, teams had a final chance to modify their group posters. Finally, the tenth session was dedicated to individual final essays students wrote, addressed to a reader new to the topic. Students did not have access to the information they used during the sessions. In these final-position essays they were able to favour either position. The ten-session workflow is presented in Fig. 1.

The length of the 10-h intervention varied from two to five weeks, depending on how the ten class hours were allocated by each teacher within the ordinary class schedule. The duration of the intervention for each teacher is presented on Table 1.

3.3. Data collection

In order to assess whether the implementation of the argument-based method by mainstream teachers is a pedagogical possibility worthy of considering, our focus was both on teachers' and students' performance. The data on which analysis was based were of two forms, students' written essays and teacher interviews, following a sequential mixed-method design (Morse, 2003).

3.3.1. Students' essays

The essay writing took place during Session 10 for the TG classes, whereas for CG classes it took place after students had spent a period of time comparable to the time the TG classes had spent on the topic. Only teacher 4 (history) asked his TG class to write the

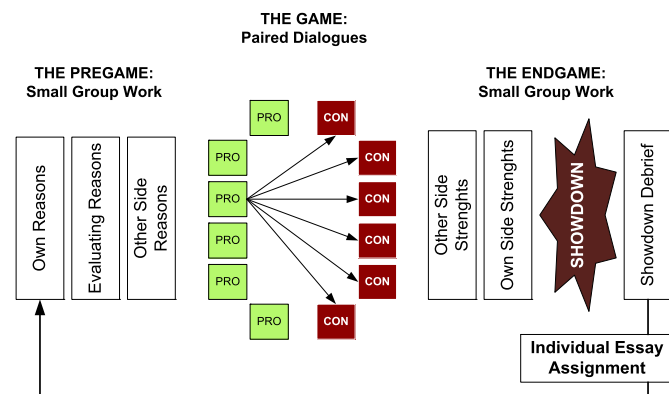


Fig. 1. The structure of the AWM curriculum.

essay on the slavery topic twice, before and after the intervention (the pre-test essay writing was not possible for the CG class, as a design constraint of the intervention was that CG students should avoid any argumentative reasoning during the intervention period, therefore the argumentative essay writing could only take place at the end of the intervention). For all other classes, the essay writing took place once, at the end of the intervention, with all TG students being previously argumentatively immersed in the topics and the CG students being informationally exposed to the topics (see section 3.1).

The prompt for the essay writing was the same for both TG and CG students, and it had the following format: *Write an opinion letter directed to the country authorities about [the concrete topic of the intervention].* This prompt was slightly changed for Teacher 4 TG and CG classes, due to the topic characteristics (slavery in the 16th century), into: *Write an opinion letter directed to a popular journal.* Across the eight classes, the only addition made for the TG students exclusively was that they could change their initial position if they wanted to (the taking of a position was not part of the CG tasks so this prompt was irrelevant for them). The time available for the essay writing was one class hour (45 min); however, the average time actually used by the students was 20 min. This was the case for both the TG and CG students.

3.3.2. Teacher interviews

The four teachers were interviewed individually once, approximately three months after the end of the intervention. The semi-structured interview, aiming at revealing their perceptions of the impact of the implementation of the AWM curriculum both on students and themselves, consisted of five questions:

1. Describe your role within the argument-based curriculum. If it was different than your traditional role, in what aspects was it different?
2. What aspects of the curriculum do you think were the most beneficial for students and why?
3. The statistical analysis showed a clear improvement in the argument quality of the written texts of the students who followed the curriculum, compared to the ones who did not.¹ How would you explain this difference?
4. Are there any other aspects in which you think your students improved, possibly as a result of the curriculum?
5. Is there anything you would do differently next time you apply the curriculum?

3.4. Data analysis

3.4.1. Students' essays

The final essays were coded using the coding scheme developed by Kuhn and colleagues (Kuhn et al., 2016b; Hemberger et al., 2017). According to this coding tool, essays are divided into idea units assigned to one of the following types: (a) units functionally supporting my side (M+), (b) units functionally weakening own side (M-), (c) units functionally supporting opposing side (O+), (d) units functionally weakening opposing side (O-), and (e) non-functional units serving none of these functions. Table 2 shows an example of each. As 'functional,' we considered any discourse unit in which an idea was supported by some type of evidence, either shared (made

available to students) or personal. Finally, adjacent idea units that take the form of opposing functions were further identified as 'However' statements (Hemberger et al., 2017; Khait, 2014). These usually are combinations of M+/M- or O+/O- or M+/O+ or M-/O- units. The importance of 'However' statements for argumentative reasoning is rooted in Billig's (1987) idea of "anti-logos" according to which every individual utterance could be matched by a counterstatement.

This coding reflects the view that a main skill manifested in argumentative discourse is the ability to look through the eyes of the other, namely, to consider counterarguments and reply to them adequately (Glassner & Schwarz, 2005; Walton, 1989). Therefore, the units coded as O+ and M-, expressing ideas that directly or indirectly support the opposing side, are considered higher-level achievements than the units coded as M+ and O-. As a consequence, units coded as O+ or M- were assigned a score of 2 points each, whereas units supporting own side received a score of 1 point each. Units that revealed non-functional statements, namely ideas that were not complete or were invalid from an argumentative point of view, received 0 points.

Inter-rater reliability calculated using Cohen's Kappa was 0.84 for the essays' coding based on all functional units described above. The inter-rater reliability score was obtained by the double coding of 30% of students' final essays, with the first coder being the author and the second coder being a researcher familiar with Kuhn et al.'s (2016a, b) coding scheme but not familiar with the data.

After each essay was coded and the final points were calculated for each, as explained above, two statistical analyses were performed: (a) the non-parametric Wilcoxon signed-rank test for non-normally distributed data (Sheskin, 2004); and (b) the parametric Student t-test, in cases where the size of the compared groups allowed for an exception to the normal distribution requirement (Gravetter & Wallnau, 2000; Stevens, 1996).

3.4.2. Teacher interviews

The four teachers' interviews were fully transcribed and translated from the original language into English. Each interview transcript was read several times, going back to the original when necessary. Although the interviewer's (author's) questions were present in the transcript, they were largely ignored during the readings to avoid biases – driven by the questions – in the classification of certain concepts within a specific topic.

The analysis of the teachers' answers followed the grounded theory approach (Corbin & Strauss, 2015; Glaser, 1978; Glaser & Strauss, 1967), applying the following steps, as summarized by Gasson (2003) and Bluff (2005):

- Open coding of transcript lines, using either *in vivo* codes (i.e. exact words used by the participants) or theoretical conceptualizations describing what participants say;
- Moving from data open coding to an axial coding through the identification of core categories;
- Using theoretical memos and questioning to capture insights on how categories are related to each other;
- Constantly comparing between words, sentences, interviews, codes and categories;
- Constructing a substantive theory by comparing the emerged categories and relations with other data and existing theories.

Moreover, my analysis of the interview transcripts was framed by a dynamic socio-constructivist and dialogic learning perspective, according to which teachers can and shall shift between different roles, each one creating different affordances for meaning making from part of the students. Students' autonomy and self-regulation is an important aspect of the learning experience,

¹ As the interview was performed after an initial analysis of the students' writing performance results, it was used as a complementary data collection process, to better understand the reasons and the extent of the positive impact of the curriculum. This mixed method approach is referred to as a sequential quantitative-qualitative design (Morse, 2003).

Table 2
Types of coded discourse units.

Code	Explanation	Example
M+	Support my side	On one side, Portugal is a very good country for receiving refugees, because after escaping from the violence they can be in a calm place without any violent act, which is what they were used to
M-	Weaken my side	but it is also true that the two welcoming houses are not enough for the expected number of refugees.
O+	Support other side	In regard to the unemployment issue, I think that in other countries they would have more job options
O-	Weaken other side	but less security
NF	Non-functional unit	for example, due to terrorist attacks in France and elsewhere.

which I was expecting to raise within a productive student-centered learning environment. Still, autonomous students can be considered “dangerous” for mainstream teachers as they, the teachers, can feel threatened by their students’ increased epistemic agency. An underlying objective of the interviews’ analysis was to examine whether this was the case for the four participant teachers in the study, and if yes, how it could be reduced. This objective did not arrive at forming a hypothesis, which would not justify the application of a grounded theory approach, because, to my knowledge, research on teachers from different disciplinary areas applying a systematic argument-base method in their classes is scarce if not absent. At least within the public Portuguese education system in which the study was carried out, such student-centered teaching practices are not common.

The reliability of the analysis in terms of trustworthiness of the categories emerged was confirmed with the synthesized member checking method (Birt, Scott, Cavers, Campbell, & Walter, 2016; Harvey, 2015). A synthesis of the final stage of analysis with the emerged themes and representative quotations from each participant for each one of the emerged themes, as well as the full transcript of the interviews, was sent by e-mail to the four teachers several months after the interviews were held. Of the four participants, one replied that she did not remember having participated in that interview, one did not reply to the e-mail, and two replied positively stating that they were highly satisfied with the analysis. None of them came back with further comments or suggestions, although this was made explicit in the e-mail as an option. In a subsequent phone conversation with one of the two participants who replied with a confirmatory member check, she anecdotally mentioned that reading her own interview transcript, without realizing that it was hers as data were anonymized to be shared with everyone, she thought that she completely agrees with what that “other participant” had said.

4. Results

4.1. Student essays

One-way independent-measures analyses of variance showed that total number of idea units (counting both functional and non-functional units) was significantly different between the CG and the TG—a median of 4.43 idea units for the CG and 8.52 for the TG, $F(1, 215) = 85.43, p < .00001$.

Final essays were analyzed according to the following criteria: (a) the type and number of functional units, namely the four types of units previously described as M+, M-, O+, and O-, and (b) the number of “However” statements appearing in the essays. Table 3 shows the frequencies of all above measures for each group.

The order of frequency among the four functional units (M+, O+, M-, O-) is the same for both the control and the treatment groups: ‘support my side’ (M+) is followed by ‘weaken other side’

Table 3
Frequencies of functional units^a and ‘However’^b statements per condition.

Functional Units	Control Group	Treatment Group
Support my side (M+)	199	356
Support other side (O+)	36	79
Weaken my side (M-)	36	93
Weaker other side (O-)	63	173
Total	334	701
‘However’ statements	22	76

^a Functional units refer to statements considered as codable because they express either an idea supporting one’s own position (M+, O-) or an idea supporting the other’s position (O+, M-).

^b ‘However’ statements are units composed of at least two functional units expressing a ‘logos-antilogos’ relation of the same idea.

(O-), which is followed by ‘weaken my side’ (M-), and ‘support other side’ (O+) (See Table 2 for an explanation of the coding categories). However, the differences across groups are notable: M+ and O+ units in the treatment group are twice the number of the ones in the control group, whereas M-, O-, and ‘However’ statements are three times more frequent in the treatment group.

In addition to these frequencies, a summary score was calculated for each final essay, assigning one (1) point for each M+ and O- unit and two (2) points for each O+ and M-unit. For the CG, the score was 3.64, with a standard deviation of 2.637 (range 0–13). For the TG it was 8.18, with a standard deviation of 4.023 (range 0–20). The essay scores of the CG and TG students were compared and a significant difference found both overall and per teacher as shown in Table 4.

Finally, only for the treatment class of Teacher 4 (TG4), who produced an essay on the ‘slavery’ topic twice, once before and once after the intervention, a significant increase in the argument quality of the total essays was found, with the pretest score of $M = 3.96$ increasing to $M = 7.59$ at posttest (Wilcoxon signed-rank test, $Z = -3.532, p < .01$). An example of how a history student improved his writing from Phase 1 to Phase 2 is shown on Table 5.

4.2. Teacher interviews

4.2.1. Functional aspects, requirements, and constraints of the AWM method in actual practice

The analysis of the teacher interviews points at specific elements of the AWM method which are responsible for bringing the observed gains to students. According to Teacher 4, the fact that the TG students got used to questioning and reaching conclusions through discussions was manifested later in their questioning behaviour in class, as well as in their attempt to identify arguments, which sometimes was also overacted:

Sometimes they exaggerate in this attempt to find arguments. Sometimes they look for where there are no arguments at all.

Table 4
Argumentation score (0–20) comparison of the final essays between the TG and the CG as a whole and per teacher.

		N	M	SD	t	gl	p
Total	TG	105	8.20	4.084	9.640	209	** .000
	CG	106	3.64	2.637			
Teacher 1	TG	24	8.92	4.262	5.322	47	** .000
	CG	25	3.76	2.260			
Teacher 2	TG	26	7.69	3.296	7.469	49	** .000
	CG	25	2.40	1.323			
Teacher 3	TG	28	8.64	4.564	2.956	56	** .005
	CG	30	5.60	3.201			
Teacher 4	TG	27	7.59	4.153	5.809	51	** .000
	CG	26	2.46	1.772			

**p < .01.

Table 5
An example of change in student writing before and after the AWM curriculum (in parenthesis the code attributed for each idea unit).

Before the intervention (History topic)	After the intervention (same student, History topic)
At the time, slavery was used as a means of labor. Slaves did not have rights neither equality. In my opinion, slavery was wrong because in that type of government it was the only means of labor (NF). From a financial point of view, slavery moved the economy and it was "good" for that time (O+), but from a social point of view slaves were very wronged and they were not even seen as humans, that said, it was horrible (O-). I think that if they abolished slavery at that time, there would be a big financial crisis (M-). To solve this without harming the economy, the best would be to give the right to choose to the slaves and to give them a salary and good work conditions (M+). My final point is that slaves should have been given the right to choose whether to work or not, and if this was not possible, the lord should at least give them some rights, good work conditions and a salary. (total score: 6).	From the Portuguese perspective, slavery was not bad, because they didn't yet have the notion of human rights (M+). However, in our actual perspective, slavery was totally unfair (M-), because killing and torturing billions for one's own benefit is totally wrong (M-). The economy of that time was sustained by slavery, and this was based on the fact that people used to practice slavery in the past (M+). Slaves were war losers and they would have been dead if it wasn't for slavery (M+). But slavery became a torture for many of them and they preferred to die than to continue being slaves (M-). In an overview, slavery is completely wrong, and they should have discovered this a long time ago. But looking at the other side, if slavery was abolished at the time, it wouldn't give the necessary means of labor for the farms and it would cause a big financial crisis. The kings and the aristocrats were forced to practice slavery with other people or to use the servitude (another type of slavery). In the times of slavery, we grew a lot in terms of science, technology and culture, perhaps if we had abolished slavery, nowadays we would have less culture and much less technology. Slavery also encouraged the commerce and the globalization (M+). However, looking through a social point of view, slavery is completely wrong, because the Portuguese took benefit of other people who were weaker and less developed (M-). (total score: 12).

But they got so much used to arguing ... They go too far. Everything is arguable. But I think it has to do with age. Being mid-school. So they think that now everything has to be ... Why Napoleon loses the war also has to be arguable (Teacher 4).

The different dialogue settings, i.e. passing from small-group discussion to peer-to-peer debate in front of the whole class, was another functional element of the AWM curriculum. According to Teacher 1, "having to talk about their opinions aloud, in a small group - in the small group and in the large group has made them orally better". Assuming the group coordinator role as part of the small group discussions further to add it to student gains, in particular to their loss of fear to talk in front of others: "(...) being in the front has become a more normal thing. Or being a group moderator is not a problem either ... this is all important" (Teacher 1).

The most important process during group discussions is, according to teachers, students' building their arguments and counterarguments together with others. This process is particularly beneficial when students are asked to defend a position contrary to their own, which was sometimes necessary because the class had to be equally distributed among the two sides. This requirement to counterargue against own ideas was shown to be highly beneficial:

I think it's very important that they are in a role they do not like. For example, if they believed in something, but have to defend the opposite, I think they improve much more, because to defend what I believe, I think it is easier. Defending what I do not believe is very difficult. And when they put themselves in the other role, I think this group ends up improving even better than those who are

defending what they believe in (Teacher 3).

Finally, some general aspects of the AWM curriculum were identified by the four teachers as positive, independently of student concrete gains. These were: the curriculum's feasibility, its age adequacy, and the fact that students' ability is not a requirement. Teachers also mentioned some conditions for a better implementation of the argument-based curriculum. One of these was the topic controversiality. Teacher 2 suggested that a different, more controversial topic than the one selected (refugees) would work even better with her 12-year-old students. Another condition for the curriculum's better implementation is teacher's modelling behaviour. This is showcased by Teacher 1 who said that: "I also had to change a bit my profile as a teacher - and this also makes them want to change their profile as students".

When it comes to constraints, curricular priorities, students' mindset, and knowledge sophistication level were the ones emerged. Regarding the curricular priorities and constraints, the AWM's structure and nature ("it involves a higher number of class hours and a more continuous work", Teacher 2) was often in conflict with the formal curriculum to fulfil which is "also extensive. And so, it is difficult to combine the two" (Teacher 3). Although, as Teacher 3 confirms, "some concepts were also soon acquired with these activities", she had to prioritize certain curricular subjects, to compensate for the class time used for the AWM project.

4.2.2. Teacher's role in the AWM curriculum

Teachers perceived their role as part of the AWM curriculum as more intelligent, more constructive, and more effective in comparison to their role implementing "as usual" teaching methods.

The reasons for that were various, but all of them relate to the socio-constructivist learning approach, according to which students are autonomous in their knowledge construction process, and the teacher functions as a knowledgeable scaffolding partner. The following concepts emerged from the analysis: teacher as a non-authority, teacher as information provider, different teaching competences, and classroom management. Below some quotes that exemplify each concept:

With the project-my profile as a teacher also grew because they grew-they suddenly began to realize that they were capable of, and therefore our profile as teachers no longer-we are not the ones who dominate all the knowledge, they are (Teacher 1).

I would give them all the information, I would tell them “check here, check over there, check the data”, but ... I think it's better if they build their own opinion (Teacher 1).

It requests from me ahh different competences-that is, competence to listen to them more, to be able to reconcile opinions, to achieve, to respect them too - I think this role is very important - to respect what they know and what they don't know yet. And therefore, it makes me ahh wiser regarding students' skills (Teacher 1).

It has transformed my way of managing the classroom. I became more interactive. I interacted more with the students (Teacher 4).

4.2.3. Students' gains as perceived by the teachers

Finally, all four teachers perceived a diversified range of skills and competences manifested by the students, in their classes but also transferred to other classes and subject matter fields. The skills identified by the teachers cover different aspects, rather than the writing skills depicted by the students' essays analysis, described in Section 4.1. The categories emerged were: cognitive skills, meta-cognitive skills, epistemological skills, communication skills, socio-emotional skills, and civic competences. The sub-categories

emerged under each one of these ‘umbrella’ concepts can be found in the grounded theory conditional matrix presented in Table 6. Here we will limit ourselves to presenting some representative quotes to showcase the diversity of skilled behaviours identified by all four teachers:

The language skills changed, there were great changes regarding self-criticism, they are more attentive to what they say, to the words they utter, to the discourse - because they have many gaps regarding the correct use of language, but nowadays they are already correcting themselves much more - so there has been a huge change not only in writing but also in orality ... and then there is a-there is a factor that it's the motivation-the self-motivation to be ... to be better (Teacher 1).

In the group there is no such fear of talking, going forward to talk to others - is much easier for these students-so it's also very good that these anxieties are disappearing quickly, and this happened a lot in this class (Teacher 1).

It was very interesting the evolution that they had throughout the year. Respect the opinion. Integrate students. Have tolerance towards students who are different, who have more difficulties. The responses are more complete. And, no longer ... It is no longer surprising for them to pose them a question that requires a more elaborate answer (Teacher 2).

I think they, as they build their own knowledge, in addition to being able to speak better, they explain themselves in a clearer way. So, writing, developing their own ideas, is also a way for most of them ... (Teacher 3).

Some students already are careful, for example, in the way they use the words ‘arguing’, ‘counter-arguing’, they use it a lot (Teacher 3).

Understand that there are other points of view. Accept better the other points of view and, in a way, learn to discuss with your points of view, other points of view. And, there, I speak of tolerance, that is, tolerate other perceptions better (Teacher 4).

Table 6

The conditional matrix emerging from the grounded theory analysis of the teacher interviews (in bold – the thematic axes, in italics – the “umbrella” categories).

When/why does it work?	Student gains	Evidence of student gains
<i>Method-related aspects:</i> Reaching conclusions through discussions	Cognitive skills: knowledge construction, acquiring processes rather than contents	<i>In-class behaviour:</i> Intellectual growth
Different dialogue settings Group coordinator role	Meta-cognitive skills: self-realisation, argument-counterargument integration	Responsibility Participation
Building arguments and counterarguments Requirement to counter-argue against own ideas	Epistemological skills: argument awareness, care for evidence	Transfer across classes/domains Text analysis/interpretation
Feasibility	Communication skills: writing and oral communication skills, building arguments and counterarguments.	Better answers and questions
Age-adequacy Ability not a requirement	Civic competences: working in group, respecting each other, constructive confrontation with others, tolerance towards other points of view	Questioning teacher Improvement in motivation and performance
<i>Other conditions/constraints:</i> Topic controversiality	Socio-emotional skills: self-image improvement, self-confidence in talking and writing, motivation	<i>Outside-class behaviour:</i> Sharing enthusiasm with friends and family Losing interest with traditional classes
Teacher modeling behaviour		
Students' mindset Curricular priorities <i>Shift in teacher's role:</i> Teacher as a non-authority Teacher as information provider Different teaching competences Classroom management		

It is also important to note that those skills appeared gradually among the TG students, remained after the intervention, and also observed by other teachers teaching different subject matters:

First of all, they start using data, because they have great difficulty when they reach the middle grades, knowing how to use information. And that was little by little and was not just for 'Argue with me'. It continued throughout the school year. That is, they have learned that they have to use data to build their knowledge, to present, to expose their way of thinking. And that was not only evident at work. It was evident throughout the school year (Teacher 2).

Today the Visual Education teacher told me that the text had not been written by students. I said "It was, it was written by the students. I have the texts they wrote". And the teacher would not believe it. So they have a way of expressing themselves in writing differently than colleagues from other classes (Teacher 2).

5. Discussion

5.1. Student gains

Any pedagogical method relying on dialogue and argumentation as its main tools lies on the principle of shifting the focus to the students rather than the teachers as the center of the instruction (Alexander, 2008; Clarke, Howley, Resnick, & Rosé, 2016; Reznitskaya & Gregory, 2013). Compared to dialogic pedagogies implementing whole-class discussions in which all students' ideas are valued and legitimized, argument-based teaching focusing on group activities aiming at peer-to-peer argumentation is even more student-centered, in the sense that the instructor is naturally obliged, due to the design of the activities, to step aside and only intervene in the groupwork when (s)he thinks necessary. Does this attitude of stepping aside work for students?

The present study provides further support for the view that students' engagement in dialogical reasoning practices has significant positive effects on their intellectual development (Alexander, 2008; Kuhn & Moore, 2015; Mercer & Dawes, 2008; Muller-Mirza & Perret-Clermont, 2009; Resnick et al., 2015). It also corroborates the findings of empirical studies explicitly showing the positive effect of dialogical argumentation on (pre)adolescents' argument writing skills (Kuhn & Crowell, 2011; Kuhn et al., 2016a; Hemberger et al., 2017; Iordanou et al., 2019; Reznitskaya et al., 2001, 2012). Most specifically, it further supports the efficacy of the AWM curriculum, which puts a premium on dense engagement with a particular topic involving both verbal and written student-to-student communication and frequent reflection. What is new, then, about our findings? In what ways and by what mechanisms does the sort of experience provided to our TG students support their intellectual development?

A first aspect that TG students appeared to develop to a greater extent than CG students was the capacity to write essays that were more argumentative, in the following ways: (a) having a larger (more than double) number of statements coded as functional, i.e., argumentative; (b) having a larger (double) number of functional statements reflecting argument construction skills, i.e., either supporting own (M+) or weakening opposing side (O-); (c) having a larger (triple) number of functional statements revealing counterargument construction skills, i.e. either weakening own (M-) or supporting opposing side (O+); and (d) having a larger (triple) number of integrated arguments (identified as 'however'

statements, Table 3). These findings indicate that although both groups were exposed to the same information regarding the topic, in their final essays TG students made use of such information as evidence to support their positions and/or weaken the opposing position, significantly more than did CG students.

This difference in argumentative writing outcomes manifested between the two groups can be attributed to the AWM curriculum for two main reasons. First, the nature of the activities held during the curriculum in its three phases (as depicted in Fig. 1) explicitly aim at the development of those argumentative reasoning skills related to the search for and use of evidence, not only as part of an individual need for cognitive support (Kim & Hannafin, 2011), but also and mainly as part of a shared, dialogic, critical thinking context with a gradually increasing sense of accountability among the participating group (Kuhn, 2019). The second reason lies in the additional finding provided by one of the classrooms (TG4), whose students wrote the individual essay on the same topic twice, once before and once after the AWM curriculum. The improvement manifested in the essays written only one month apart was evident: The structure of the texts was more argumentative, including more counterarguments as well as integrated arguments, both of them key components of argumentative reasoning and writing (Billig, 1987; Nussbaum, 2008; Nussbaum & Schraw, 2007).

How can these achievements be explained? The predominant view is that through their engagement in dialogical argumentation practices, students acquire a new 'argument schema,' meaning a synergy of declarative, procedural, and conditional knowledge of what an argument is, and how and when to use it (Reznitskaya & Anderson, 2002; Reznitskaya et al., 2009, 2012). According to this view, students acquire more advanced discursive strategies from an argumentation point of view, as a result of their dialogical practice, which they subsequently transfer to writing (Anderson et al., 1997, 2001; Reznitskaya et al., 2001, 2009). However, research is still divided in regards to what type of dialogical practice is more efficient for argument skills to be developed orally, and subsequently transferred to writing.

Concerning peer-to-peer dialogue and its benefits, the literature is controversial. On one hand are studies that show that dialogue with adults is more beneficial than dialogue with peers (Coirier, Coquin-Viennot, Golder, & Passerault, 1990; Grossen, Liengme Bessire, & Perret-Clermont, 1997). On the other hand, extensive research has shown the great potential in classroom contexts of peer-to-peer dialogue in dyads or small groups, compared to teacher-led whole-class interactions (e.g. Crowell & Kuhn, 2014; Hogan, Nastasi, & Pressley, 1999; Kuhn & Moore, 2015). In addition, several studies suggest that it is the practice of argumentative language with same-level peers that affects the way adolescents individually think, elaborate, and revise the ideas discussed (Kuhn, 2018a, 2019; Larrain, 2017; Larrain, Freire, López, & Grau, 2019). Therefore, one factor contributing to the improvement of TG students' writing performance in the present study could be the significant amount of time arguing and counterarguing that they spent with each other, and in different formats (small groups, dyads, dyads vs dyads, and one-to-one), as also emerged from the teacher interviews. In engaging in rich peer-to-peer dialogue, students naturally use discourse argument strategies, which develop earlier and with greater ease than do written ones (Ferreti & Graham, 2019; Kuhn, 2019). Exercise of these strategies strengthens them. In addition, the argument-related cognitive and metacognitive resources noted above are activated, and accordingly put to use in individual writing.

We can also draw on the fact that teachers perceived a series of changes in their TG students, across cognitive, metacognitive, epistemological, and social-emotional skills. Students not only

changed in the way they wrote essays, which was the final explicit outcome of the AWM curriculum; they also were reported to ask more questions during subsequent classes, were attentive to evidence and its use, became more expressive in their stances, and felt more confident to communicate either with the teacher and most important with their peers. How do these different behaviors relate to each other and what do they have to do with argumentation? An answer may lie in the immersion approach and its pedagogical potential, rather than the argument curriculum *per se*. According to Prawat (1991), the immersion approach puts “its emphasis on ideas as the most important resource in promoting thought” (p. 6). Students’ understanding of what an idea is, what it means to have an idea, or to promote it through valid arguments may be an important gain yielded by an immersion approach to argumentation.

5.2. Teacher gains

Dialogue-based teaching typically consists of either whole-class discussions exclusively, or a mix of whole-class and small-group discussions (Mercer & Dawes, 2008; Resnick et al., 2015; Reznitskaya & Gregory, 2013). The problem with teachers moderating whole-class discussions is the difficulty they experience in becoming confident, authentic dialogic teachers, disposed to facilitate and mediate students’ contributions with the aim of reinforcing their participation and accountability (Kuhn, 2019; Michaels, O’Connor, & Resnick, 2008). This problem is not as severe when dialogical argumentation is centered *among* students, rather than continuously mediated by teachers. However, placing the focus on peer-to-peer argumentation, and in different formats, as is the case in the AWM curriculum, requires a different type of coaching by teachers, different than that involved in mediating whole-class discussions. Teachers are required to act as ‘reasoning gatekeepers’: they need to ensure that students reason in ways consistent with the particular goals of the activity, for example, by introducing specific teacher moves that help student-student discourse to be more productive (Sandoval, Kwako, Modrek, & Kawasaki, 2018). Teacher modeling behaviour emerged as an important aspect also in the teacher interviews. Further analysis of teachers’ attitudes and perceptions towards argument-based teaching in the whole-class versus student-to-student conditions would shed light on the distinct challenges of both approaches and how each are best met.

By placing the focus on students and their mutual exchanges, another type of authority comes to replace the epistemic authority traditionally bestowed on the teacher: the authority of regulating the distribution of epistemic agency among students or groups of students (Miller, Manz, Russ, Stroupe, & Berland, 2018), and of orchestrating one or more activities taking place in a classroom (Dillenbourg, 2013). This twofold authority of epistemic and classroom discourse and/or activity regulation is a crucial aspect of facility in argument-based pedagogical practice, known to be difficult to acquire (Clarke et al., 2016; Sedova et al., 2016; Wilkinson et al., 2017). However, by focusing on enactment of the structured components of the AWM curriculum, teachers in our study appeared to acquire awareness of being epistemic regulators – namely the ones who facilitate or assist knowledge construction, and not the ones who impose it – as their interviews showed. Although we did not seek to compare here the whole-class versus AWM student-to-student focused approach, the reported and observed ease with which the four participant teachers implemented the AWM curriculum can be contrasted with the frequently reported difficulty regarding dialogic teachers’ management of whole-class discussions (Clarke et al., 2016).

The feasibility of dialogical argumentation as a pedagogical method in this study can be partially attributed to the fact that the

four participant teachers, in addition to training by modeling, were also given the lesson plans with the exact activities to do for each session (see Appendix 1). Their only preparation was to carefully read the planning of the activities beforehand, organize the necessary materials (post-it notes, papers, pencils, etc.), and take care of the time. Of the two main challenges for 21st century teachers described as being a coach and an instructional designer (Hoogveld, Paas, & Jochems, 2005), the teachers in our study only had to address the former, as the AWM method was already adapted to their curricular contents and interest by the research team. A future study shall include both aspects, designing the activities and coaching them, as part of the teacher professional development, to see whether further benefits emerge both for teachers and students alike.

The same advantage described above, i.e. the fact that the lesson plans were constructed by the research team and not by the teachers, is also a study’s limitation. This limitation is important if we want to consider a more generalized implementation of the AWM curriculum outside a teacher training context. The constructed lesson plans mainly functioned as an outline of the student-centered activities, rather than pedagogical guidelines of how teachers should behave in activities they themselves define. Having an outline of the sequence of the activities and their approximate timing is an important aspect of any type of instructional design intervention (Weinberger, Stegmann, & Fischer, 2010). However, when it comes to the actual contents/activities of the curriculum, these, in the case of AWM, do not have to be designed anew every time, which is what teachers are usually hesitant about (Reiser et al., 2000). Adoption of the AWM curriculum merely requires choice of the topic, according to the teacher’s/ students’ interests, and organization of the activities per class. This simplicity is what enables the transdisciplinary, cross-curricular potential, discussed subsequently.

5.3. Dialogical argumentation as a transdisciplinary pedagogical method

In our study, argumentative reasoning was significantly better for the TG students as compared to the CG students for all four different subject matter classes. This finding opens the discussion about how general argumentative reasoning skills are. Although not exclusively referring to argumentative reasoning, Perkins and Salomon (1989) argue that “[T]here are general cognitive skills; but they always function in contextualized ways” (p. 19). They bring the example of academic philosophers, who tend to show the general cognitive strategy of using counterexamples to test claims. Rather than examining the context-boundedness of this strategy (e.g. *Will it be that they can only construct counterexamples in domains where they have a good knowledge base?*), they opt for defining the characteristics of “a general, learnable and worthwhile” (Perkins & Salomon, 1989, p. 19) cognitive skill. These are: (a) that it shows seeming use, which means that it may appear that it is general, while it is entirely contextualized; (b) it plays an important role, given that general skills often serve as catalysts for the development of subject-specific thinking skills (Higgins & Baumfield, 1998); (c) it is transferrable, in the sense of promoting learning and expertise in different fields (Higgins & Baumfield, 1998); and (d) it is commonly absent, meaning that a skill is differentiated by a generally common ability.

Based on the criteria above, we could argue that argumentative reasoning is a set of general skills (i.e. constructing arguments using evidence, constructing and defeating counter-arguments, etc.), applicable in different domains. There is sufficient evidence to claim that this is true. For example, Topcu, Sadler, and Yilmaz-Tuzun (2010) suggested that socioscientific reasoning is not

strongly issue-dependent and it can be more ‘transferable’ across distinct contexts. Similarly, several studies applying the AWM curriculum showed transfer of argument skills across topics within the same domain (e.g. [Iordanou & Kuhn, 2020](#); [Kuhn, Goh, Iordanou, & Shaenfield, 2008](#)), or even across domains ([Iordanou, 2010](#)). However, the problem of transdisciplinarity arises when disciplinary reasoning becomes a part, or even requisite, of argumentative reasoning. This is the case where specific domain knowledge is necessary for an argument to be sound, which is not the case when arguing about general and social issues ([Means & Voss, 1996](#)). Toulmin (1958) first argued that although different domains share many elements of argument structure and reasoning, they also have different criteria for what counts as evidence. Drawing on this, several educational researchers (e.g. [De La Paz et al., 2012](#); [Windschitl, Thompson, & Braaten, 2008](#)) claim that for students to be able to reason as scientists, historians, experts, they need to know how to argue in that particular domain, as the rules of inference are different from discipline to discipline ([Becher, 1994](#)). Where does the truth lie? Is argumentative reasoning context-dependent or context-independent?

Our answer, supported by our findings, is that argumentative reasoning is context-independent *and* content-bound. It is context-independent, as the same method, as depicted in a curriculum aiming at general argument reasoning skills, worked across different disciplinary domains. At the same time, it is content bound, as for students to be able to argue about the given topics, information was made available to them in form of defeasible knowledge (QAs). This dialogic representation of knowledge situated information, and subsequently evidence, as part of an inquiry process in which “evidence is marshalled to support or defeat claims to knowledge” ([Walton, 2011](#), p. 131). Moreover, the richness of dialogic interaction as part of the AWM curriculum heightens students’ ability to see and practice the epistemological power and value of evidence, which lies in the heart of the cross-disciplinary value of argumentation. As [Litman and Greenleaf \(2018\)](#) conclude, “interactive argumentation is in itself a context for building students’ understanding of content, as well as argumentation” (p. 15). Our study opens the path for argument-based pedagogical methods to be introduced in different subject matters across the curriculum.

6. Conclusion

In conclusion, this study showed that it is possible for mainstream teachers to implement dialogical argumentation as a student-centered pedagogical method, with only a 12-h teacher training and an outlined structure of lesson plans. The four participant teachers perceived their role as a coach within the AWM dialogical method as more intelligent, more constructive, and more effective in comparison to their role implementing “as usual” teaching methods, even when the latter includes whole-class and/or small-group discussion, yet without any intentional engagement in argumentative objectives or structure.

Teachers’ use of the method was proven to be an effective pedagogical tool in terms of argumentative reasoning gains, manifested in students’ writing. This was the case across the four subject areas of the study, namely language arts, civic education, science and history, as well as across core argumentative reasoning skills, manifested in the use of functional evidence-based claims as building blocks of written arguments. In addition, students who participated in the 10-h AWM curriculum manifested a range of cognitive, metacognitive, epistemological, civic, and socio-emotional skills, as emerged from the analysis of the teacher interviews. This finding opens the path for more teacher-researcher collaborations towards the understanding and implementation of

21st century critical literacy skills.

A limitation of the study is the fact that the teacher interviews that were used to better interpret the student performance findings were only held once and with only four teachers. Although the analysis offered us insights into how and why students’ and teachers’ gains were observed, as well as some constraints, we cannot yet form a complete grounded theory about how argument-based curricula work from a teachers’ perspective and what it takes for teachers to give up their traditional role and adopt the one of a coach, who mediates rather than guides students’ autonomous knowledge construction. Future research will give a closer look at this need for a bottom-up insight on how teachers interact with student-centered argument curricula, and what researchers can do to further empower them.

Funding

This work is funded by national funds through the FCT – Fundação para a Ciência e a Tecnologia, I.P., through the project METACARE (Project code: PTDC/FER-FIL/28278/2017) and under the Norma Transitória (contract number: DL 57/2016/CP1453/CT0066).

Declaration of competing interest

None.

Acknowledgements

The author would like to thank Filipa Barreto, Pedro Mil-Homens, Octávio Gameiro, Paula Pereira and Maria José Vilas Boas, and their students, for their enthusiastic participation in the project, as well as three anonymous reviewers for their constructive comments that helped improve the manuscript.

APPENDIX 1

Example of a lesson sequence plan as part of the Pre-Game.

- Students sit in already made groups of 4–6. It would be good if students in favour of A sit separately than students in favour of B. The teacher starts by giving some explanations regarding what they would be doing in the following weeks (suggested time: 7 min).
- Some of the things that the teacher can explain as an Introduction to the AWM classes:
 - Argumentation is like a volleyball game. We have two sides, and we need to make sure that the ball passes from side to side.
 - This ball represents our reasons and evidence. We support our reasons with evidence; the stronger our reasons are, the stronger the “hit”.
 - At the end, the winning side is the one that manages best to support their opinions with reasons and evidence.
- After that introduction, the teacher explains what the first task will be: it will be about working in same-side groups with the aim of generating at least three main reasons that support our side. Every group has an A3 poster and several Post-it notes of the same colour. The teacher may also ask the groups to nominate a leader for this task, which is composed of the following sub-tasks (the teacher explains one a time giving appropriate time for each):
 - a) Each student writes individually which is his/her main reason to support Side A or B. The students use Post-it notes

- given to them by the group leader or the teacher (suggested time: 3 min).
- Each student turns to the colleague on his/her right side and together they talk about their individual reasons, check if they are the same or similar, whether they are clear enough, etc. If their reasons match, they rewrite them in one Post-it and in a more synthetic way (suggested time: 3 min).
 - Now the whole group with the help of the leader compares the reasons produced by the members of the group to each other and repeat the previous step: they put together the ideas that are similar and re-write in a synthetic way all the non-duplicated reasons. (suggested time: 5 min).
 - Each group evaluates their reasons and decides on a ranking from the most to the least important. The Post-it notes with the rewritten reasons are collocated on the poster in the decided order. (suggested time: 5 min).
 - In Post-it notes of a different colour, each student individually writes two questions asking for information that (s)he thinks necessary to support two different reasons of the ones that are on the group poster. (suggested time: 5 min).
 - Each student shows their questions to the colleague on their right side, with whom they compare and rewrite their questions. At the end the whole group writes a list of at least three questions about things they would like to have information so that they can reinforce their reasons. (suggested time: 12 min).

APPENDIX 2

Treatment fidelity checklist.

- CG students participated in whole-class discussions dealing with the given topic.
- CG students were encouraged to formulate their arguments and/or counterarguments during those discussions.
- CG students were encouraged to express their opinions about the given topic during those discussions.
- CG students negotiated ideas during small group discussions.
- CG students were exposed to all the information available for the topic.
- CG students were challenged and often asked to justify and defend their opinions.
- CG students were encouraged to interact with each other through engaging in constructive dialogue.
- CG students were engaged in an individual or group task related to the topic that did not include dialogue and/or argumentation about the dilemma of the EG.
- CG students spent an amount of time working with the given topic and information that is comparable to the EG dedication of time.
- The time assigned for the CG final essays was comparable (more or less the same) to the time assigned for the EG final essays.

Likert-scale questions (repeated for both the CG and TG classes):

- How often did students directly address another student, by adding on to what said?
 - Rarely; b) sometimes; c) quite often; d) often; e) very often
- How often did students directly address another student, by challenging what said?
 - Rarely; b) sometimes; c) quite often; d) often; e) very often
- How often did students directly address the teacher, by adding on to what said?
 - Rarely; b) sometimes; c) quite often; d) often; e) very often

- Rarely; b) sometimes; c) quite often; d) often; e) very often
- How often did students directly address the teacher, by challenging what said?
 - Rarely; b) sometimes; c) quite often; d) often; e) very often
 - How often did students listen w/out speaking?
 - Rarely; b) sometimes; c) quite often; d) often; e) very often

References

- Alexander, R. J. (2008). *Towards dialogic teaching: Rethinking classroom talk* (4th ed.). York, UK: Dialogos.
- Anderson, R. C., Chinn, C., Chang, J., Waggoner, M., & Yi, H. (1997). On the logical integrity of children's arguments. *Cognition and Instruction*, 15(2), 135–167. https://doi.org/10.1207/s1532690xci1502_1
- Anderson, R. C., Nguyen-Jahiel, K., McNurlen, B., Archodidou, A., Kim, S., Reznitskaya, A., et al. (2001). The snowball phenomenon: Spread of ways of talking and ways of thinking across groups of children. *Cognition and Instruction*, 19(1), 1–46. https://doi.org/10.1207/s1532690xci1901_1
- Becher, T. (1994). The significance of disciplinary differences. *Studies in Higher Education*, 19(2), 151–161.
- Berland, L. K., & Reiser, B. J. (2009). Making sense of argumentation and explanation. *Science Education*, 93(1), 26–55. <https://doi.org/10.1002/sce.20286>
- Billig, M. (1987). *Arguing and thinking: A rhetorical approach to social psychology*. Cambridge, UK: Cambridge University Press.
- Birt, L., Scott, S., Cavers, D., Campbell, C., & Walter, F. (2016). Member checking: A tool to enhance trustworthiness or merely a nod to validation? *Qualitative Health Research*, 26(13), 1802–1811. <https://doi.org/10.1177/1049732316654870>
- Bluff, R. (2005). Grounded theory: The methodology. In I. Holloway (Ed.), *Qualitative research in health care* (pp. 147–167). London: Open University Press.
- Brown, M. W. (2009). The teacher-tool relationship: Theorizing the design and use of curriculum materials. In J. Remillard, B. Herbel-Sisenham, & G. Lloyd (Eds.), *Mathematics teachers at work: Connecting curriculum materials and classroom instruction* (pp. 17–37). New York, NY: Routledge.
- Chen, Y. C., Hand, B., & Park, S. (2016). Examining elementary students' development of oral and written argumentation practices through argument-based inquiry. *Science & Education*, 25(3–4), 277–320. <https://doi.org/10.1007/s11191-016-9811-0>
- Chinn, C. A., & Brewer, W. F. (1998). An empirical test of a taxonomy of responses to anomalous data in science. *Journal of Research in Science Teaching*, 35(6), 623–654. [https://doi.org/10.1002/\(sici\)1098-2736\(199808\)35:6<623::aid-tea3>3.0.co;2-o](https://doi.org/10.1002/(sici)1098-2736(199808)35:6<623::aid-tea3>3.0.co;2-o)
- Christoph, J. N., & Nystrand, M. (2001). Taking risks, negotiating relationships: One teacher's transition toward a dialogic classroom. *Research in the Teaching of English*, 249–286.
- Clarke, S. N., Howley, I., Resnick, L., & Rosé, C. P. (2016). Student agency to participate in dialogic science discussions. *Learning, Culture and Social Interaction*, 10, 27–39. <https://doi.org/10.1016/j.lcsi.2016.01.002>
- Coirier, P., Coquin-Viennot, D., Golder, D., & Passerault, J. M. (1990). Le traitement du discours argumentatif: Recherches en production et en compréhension. *Archives de Psychologie*, 58, 315–348.
- Corbin, J., & Strauss, A. L. (2015). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (4th ed.). Thousand Oaks, CA: Sage.
- Crowell, A., & Kuhn, D. (2014). Developing dialogic argumentation skills: A 3-year intervention study. *Journal of Cognition and Development*, 15(2), 363–381. <https://doi.org/10.1080/15248372.2012.725187>
- De La Paz, S., & Felton, M. K. (2010). Reading and writing from multiple source documents in history: Effects of strategy instruction with low to average high school writers. *Contemporary Educational Psychology*, 35(3), 174–192. <https://doi.org/10.1016/j.cedpsych.2010.03.001>
- De La Paz, S., Ferretti, R., Wissinger, D., Yee, L., & MacArthur, C. (2012). Adolescents' disciplinary use of evidence, argumentative strategies, and organizational structure in writing about historical controversies. *Written Communication*, 29(4), 412–454. <https://doi.org/10.1177/0741088312461591>
- Dillenbourg, P. (2013). Design for classroom orchestration. *Computers & Education*, 69, 485–492. <https://doi.org/10.1016/j.compedu.2013.04.013>
- Duschl, R. A., & Osborne, J. (2002). Supporting and promoting argumentation discourse in science education. *Studies in Science Education*, 38, 39–72. <https://doi.org/10.1080/03057260208560187>
- Erduran, S., Simon, S., & Osborne, J. (2004). TAPping into argumentation: Developments in the application of Toulmin's Argument Pattern for studying science discourse. *Science Education*, 88, 915–933. <https://doi.org/10.1002/sce.20012>
- Evagorou, M., & Dillon, J. (2011). Argumentation in the teaching of science. In D. Corrigan, et al. (Eds.), *The professional knowledge base of science teaching* (pp. 189–203). Dordrecht, Netherlands: Springer.
- Evagorou, M., Jimenez-Aleixandre, M. P., & Osborne, J. (2012). 'Should we kill the grey squirrels?' A study exploring students' justifications and decision-making. *International Journal of Science Education*, 34(3), 401–428. <https://doi.org/10.1080/09500693.2011.619211>
- Ferretti, R. P., & Graham, S. (2019). Argumentative writing: Theory, assessment, and instruction. *Reading and Writing*, 32, 1345–1357. <https://doi.org/10.1007/s11145->

- 019-09950-x
- Ferretti, R. P., & Lewis, W. E. (2013). Best practices in teaching argumentative writing. In S. Graham, C. A. MacArthur, & J. Fitzgerland (Eds.), *Best practices in writing instruction* (2nd ed., pp. 113–140). New York & London: Guilford Press.
- Gasson, S. (2003). Rigor in grounded theory research: An interpretive perspective on generating theory from qualitative field studies. In M. Whitman, & A. Woszczyński (Eds.), *The handbook of information systems research* (pp. 79–102). Hershey, PA: IGI Global.
- Glaser, B. G. (1978). *Advances in the methodology of grounded theory*. Mill Valley, CA: Sociology Press.
- Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. New Brunswick, NJ: AldineTransaction.
- Glassner, A., & Schwarz, B. B. (2005). The antilogos ability to evaluate information supporting arguments. *Learning and Instruction*, 15(4), 353–375. <https://doi.org/10.1016/j.learninstruc.2005.07.002>
- Gravetter, F. J., & Wallnau, L. B. (2000). *Statistics for the behavioral sciences* (5th ed.). Belmont, CA: Wadsworth.
- Grossen, M., Liengme Bessire, M.-J., & Perret-Clermont, A.-N. (1997). Construction de l'interaction et dynamiques sociocognitives. In M. Grossen, & B. Py (Eds.), *Pratiques sociales et médiations symboliques* (pp. 221–247). Bern, Switzerland: Peter Lang.
- Harvey, L. (2015). Beyond member-checking: A dialogic approach to the research interview. *International Journal of Research and Method in Education*, 38(1), 23–38. <https://doi.org/10.1080/1743727X.2014.914487>
- Hemberger, L., Kuhn, D., Matos, F., & Shi, Y. (2017). A dialogic path to evidence-based argumentative writing. *The Journal of the Learning Sciences*, 26(4), 575–607. <https://doi.org/10.1080/10508406.2017.1336714>
- Herrenkohl, L. R., & Cornelius, L. (2013). Investigating elementary students' scientific and historical argumentation. *The Journal of the Learning Sciences*, 22(3), 413–461. <https://doi.org/10.1080/10508406.2013.799475>
- Higgins, S., & Baumfield, V. (1998). A defence of teaching general thinking skills. *Journal of Philosophy of Education*, 32(3), 391–398. <https://doi.org/10.1111/1467-9752.00103>
- Hogan, K., Nastasi, B. K., & Pressley, M. (1999). Discourse patterns and collaborative scientific reasoning in peer and teacher-guided discussions. *Cognition and Instruction*, 17(4), 379–432. https://doi.org/10.1207/S1532690XC1704_2
- Hoogveld, A. W., Paas, F., & Jochems, W. M. (2005). Training higher education teachers for instructional design of competency-based education: Product-oriented versus process-oriented worked examples. *Teaching and Teacher Education*, 21(3), 287–297. <https://doi.org/10.1016/j.tate.2005.01.002>
- Iordanou, K. (2010). Developing argument skills across scientific and social domains. *Journal of Cognition and Development*, 11(3), 293–327. <https://doi.org/10.1080/15248372.2010.485335>
- 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., Matos, F., Shi, Y., & Hemberger, L. (2019). *Learning by arguing. Learning and instruction* (Vol. 63). <https://doi.org/10.1016/j.learninstruc.2019.05.004>. Online first.
- Iordanou, K., & Rapanta, C. (2021). Argue with me: A method for developing argument skills. *Frontiers in Psychology*, 12. <https://doi.org/10.3389/fpsyg.2021.631203>
- Khait, V. (2014). *Making use of the dual functions of evidence in adolescents' argumentation*. Unpublished Phd Dissertation, Teachers' College. New York: Columbia University.
- Kim, M. C., & Hannafin, M. J. (2011). Scaffolding problem solving in technology-enhanced learning environments (TELEs): Bridging research and theory with practice. *Computers & Education*, 56(2), 403–417. <https://doi.org/10.1016/j.compedu.2010.08.024>
- Kuhn, D. (1991). *The skills of argument*. New York, NY: Cambridge University Press.
- Kuhn, D. (2018a). A role for reasoning in a dialogic approach to critical thinking. *Topoi*, 37(1), 121–128. <https://doi.org/10.1007/s11245-016-9373-4>
- Kuhn, D. (2018b). *Building our best future: Thinking critically about ourselves and our world*. New York, NY: Wessex.
- Kuhn, D. (2019). Critical thinking as discourse. *Human Development*, 62(3), 146–164. <https://doi.org/10.1159/000500171>
- Kuhn, D., & Crowell, A. (2011). Dialogic argumentation as a vehicle for developing young adolescents' thinking. *Psychological Science*, 22(4), 545–552. <https://doi.org/10.1177/0956797611402512>
- Kuhn, D., Feliciano, N., & Kostikina, D. (2019). Engaging contemporary issues as practice for citizenship. *The Social Studies*, 1–13. <https://doi.org/10.1080/00377996.2019.1625856>
- 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, 1311–1329. <https://doi.org/10.1111/j.1467-8624.2008.01190.x>
- Kuhn, D., Hemberger, L., & Khait, V. (2016a). 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., Hemberger, L., & Khait, V. (2016b). *Argue with me: Developing thinking and writing through dialog*. New York: Routledge.
- Kuhn, D., & Moore, W. (2015). Argument as core curriculum. *Learning: Research and Practice*, 1, 66–78. <https://doi.org/10.1080/23735082.2015.994254>
- Kuhn, D., Zillmer, N., Crowell, A., & Zavala, J. (2013). Developing norms of argumentation: Metacognitive, epistemological, and social dimensions of developing argumentative competence. *Cognition and Instruction*, 31(4), 456–496. <https://doi.org/10.1080/07370008.2013.830618>
- Larrain, A. (2017). Group-work discussions and content knowledge gains: Argumentative inner speech as the missing link? *Learning, Culture and Social Interaction*, 14, 67–78. <https://doi.org/10.1016/j.lcsi.2017.04.002>
- Larrain, A., Freire, P., & Howe, C. (2014). Science teaching and argumentation: One-sided versus dialectical argumentation in Chilean middle-school science lessons. *International Journal of Science Education*, 36(6), 1017–1036. <https://doi.org/10.1080/09500693.2013.832005>
- Larrain, A., Freire, P., López, P., & Grau, V. (2019). Counter-arguing during curriculum-supported peer interaction facilitates middle-school students' science content knowledge. *Cognition and Instruction*. <https://doi.org/10.1080/07370008.2019.1627360>. Online first.
- Larrain, A., Moreno, C., Grau, V., Freire, P., Salvat, I., López, P., et al. (2017). Curriculum materials support teachers in the promotion of argumentation in science teaching: A case study. *Teaching and Teacher Education*, 67, 522–537.
- Lieberman, A. (1995). Practices that support teacher development: Transforming conceptions of professional learning. In F. Stevens (Ed.), *Innovating and evaluating science education* (pp. 67–78). National Science Foundation.
- Litman, C., & Greenleaf, C. (2018). Argumentation tasks in secondary English language arts, history, and science: Variations in instructional focus and inquiry space. *Reading Research Quarterly*, 53(1), 107–126. <https://doi.org/10.1002/rq.187>
- Martin, A. M., & Hand, B. (2009). Factors affecting the implementation of argument in the elementary science classroom. A longitudinal case study. *Research in Science Education*, 39(1), 17–38. <https://doi.org/10.1007/s11165-007-9072-7>
- McNeill, K. L. (2009). Teachers' use of curriculum to support students in writing scientific arguments to explain phenomena. *Science Education*, 93(2), 233–268. <https://doi.org/10.1002/sce.20294>
- McNeill, K. L., González-Howard, M., Katsh-Singer, R., & Loper, S. (2017). Moving beyond pseudoargumentation: Teachers' enactments of an educative science curriculum focused on argumentation. *Science Education*, 101(3), 426–457. <https://doi.org/10.1002/sce.21274>
- McNeill, K. L., & Pimentel, D. S. (2010). Scientific discourse in three urban classrooms: The role of the teacher in engaging high school students in argumentation. *Science Education*, 94(2), 203–229. <https://doi.org/10.1002/sce.20364>
- Means, M. L., & Voss, J. F. (1996). Who reasons well? Two studies of informal reasoning among children of different grade, ability, and knowledge levels. *Cognition and Instruction*, 14(2), 139–178. https://doi.org/10.1207/s1532690xc1402_1
- Mercer, N., & Dawes, L. (2008). The value of exploratory talk. In N. Mercer, & S. Hodgkinson (Eds.), *Exploring talk in school* (pp. 55–71). London, UK: Sage.
- Michaels, S., O'Connor, C., & Resnick, L. B. (2008). Deliberative discourse idealized and realized: Accountable talk in the classroom and in civic life. *Studies in Philosophy and Education*, 27(4), 283–297. <https://doi.org/10.1007/s11217-007-9071-1>
- Miller, E., Manz, E., Russ, R., Stroupe, D., & Berland, L. (2018). Addressing the epistemic elephant in the room: Epistemic agency and the next generation science standards. *Journal of Research in Science Teaching*, 55(7), 1053–1075. <https://doi.org/10.1002/tea.21459>
- Morse, J. M. (2003). Principles of mixed methods and multimethod research design. In A. Tashakkori, & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral sciences* (189–208). London, UK: Sage.
- Muller-Mirza, N., & Perret-Clermont, A. N. (2009). *Argumentation and education: Theoretical foundations and practices*. Dordrecht, Netherlands: Springer.
- Newell, G. E., Beach, R., Smith, J., & VanDerHeide, J. (2011). Teaching and learning argumentative reading and writing: A review of research. *Reading Research Quarterly*, 46(3), 273–304. <https://doi.org/10.1598/RRQ.46.3.4>
- Nussbaum, E. M. (2008). Collaborative discourse, argumentation, and learning: Preface and literature review. *Contemporary Educational Psychology*, 33(3), 345–359. <https://doi.org/10.1016/j.cedpsych.2008.06.001>
- Nussbaum, E. M., & Schraw, G. (2007). Promoting argument-counterargument integration in students' writing. *The Journal of Experimental Education*, 76(1), 59–92. <https://doi.org/10.3200/JEXE.76.1.59-92>
- Osborne, J., Erduran, S., & Simon, S. (2004). *Ideas, evidence and argument in science (IDEAS) project*. London, UK: Nuffield Foundation.
- Perkins, D. N., & Salomon, G. (1989). Are cognitive skills context-bound? *Educational Researcher*, 18(1), 16–25.
- Pimentel, D. S., & McNeill, K. L. (2013). Conducting talk in secondary science classrooms: Investigating instructional moves and teachers' beliefs. *Science Education*, 97(3), 367–394. <https://doi.org/10.1002/sce.21061>
- Prawat, R. S. (1991). The value of ideas: The immersion approach to the development of thinking. *Educational Researcher*, 20(2), 3–30. <https://doi.org/10.3102/0013189X020002003>
- Rapanta, C. (2019). *Argumentation strategies in the classroom*. Delaware: Vernon Press.
- Rapanta, C., & Macagno, F. (2019). Pragmatics, education, and argumentation: Introduction to the Special Issue. *Learning, Culture & Social Interaction*. <https://doi.org/10.1016/j.lcsi.2019.100371>
- Rapanta, C., & Tróvão, S. (2020). Shall we receive more refugees or not? A comparative analysis and assessment of Portuguese adolescents' arguments and social representations. *Pedagogy, Culture & Society*, 28(4), 581–603.
- Reiser, B. J., Spillane, J. P., Steinmuller, F., Sorsa, D., Carney, K., & Kyza, E. (2000). Investigating the mutual adaptation process in teachers' design of technology-infused curricula. In B. Fishman, & S. O' Connor-Divelbiss (Eds.), *Fourth*

- international conference of the learning sciences (pp. 342–349). Mahwah, NJ: Erlbaum.
- Resnick, L. B., Asterhan, C. S. C., & Clarke, S. N. (Eds.). (2015). *Socializing intelligence through academic talk and dialogue*. Washington, DC: American Educational Research Association.
- Reznitskaya, A., Anderson, R. C., McNurlen, B., Nguyen-Jahiel, K., Archodidou, A., & Kim, S. Y. (2001). Influence of oral discussion on written argument. *Discourse Processes*, 32(2–3), 155–175. https://doi.org/10.1207/S15326950DP3202&3_04
- Reznitskaya, A., Glina, M., Carolan, B., Michaud, O., Rogers, J., & Sequeira, L. (2012). Examining transfer effects from dialogic discussions to new tasks and contexts. *Contemporary Educational Psychology*, 37(4), 288–306. <https://doi.org/10.1016/j.cedpsych.2012.02.003>
- Reznitskaya, A., & Gregory, M. (2013). Student thought and classroom language: Examining the mechanisms of change in dialogic teaching. *Educational Psychologist*, 48(2), 114–133. <https://doi.org/10.1080/00461520.2013.775898>
- Reznitskaya, A., Kuo, L. J., Clark, A. M., Miller, B., Jadallah, M., Anderson, R. C., et al. (2009). Collaborative reasoning: A dialogic approach to group discussions. *Cambridge Journal of Education*, 39(1), 29–48. <https://doi.org/10.1080/03057640802701952>
- Sampson, V., & Blanchard, M. R. (2012). Science teachers and scientific argumentation: Trends in view and practice. *Journal of Research in Science Teaching*, 49(9), 1122–1148. <https://doi.org/10.1002/tea.21037>
- Sampson, V., Grooms, J., & Walker, J. P. (2011). Argument-driven inquiry as a way to help students learn how to participate in scientific argumentation and craft written arguments: An exploratory study. *Science Education*, 95(2), 217–257. <https://doi.org/10.1002/sce.20421>
- Sandoval, W., Kwako, A. J., Modrek, A., & Kawasaki, J. (2018). Patterns of classroom talk through participation in discourse-focused teacher professional development. In *Proceedings of the 13th annual international conference of the learning sciences (Vol. 2)* (London, UK).
- Sedova, K., Sedlacek, M., & Svaricek, R. (2016). Teacher professional development as a means of transforming student classroom talk. *Teaching and Teacher Education*, 57, 14–25. <https://doi.org/10.1016/j.tate.2016.03.005>
- Shank, R. (2011). *Teaching minds: How cognitive science can save our schools*. New York, NY: Teachers College Press.
- Sheskin, D. (2004). *Handbook of parametric and nonparametric statistical procedures* (3rd ed.). Boca Raton: Chapman & Hall/CRC.
- Stevens, J. (1996). *Applied multivariate statistics for the social sciences* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum.
- Topcu, M. S., Sadler, T. D., & Yilmaz-Tuzun, O. (2010). Preservice science teachers' informal reasoning about socioscientific issues: The influence of issue context. *International Journal of Science Education*, 32(18), 2475–2495. <https://doi.org/10.1080/09500690903524779>
- Venville, G. J., & Dawson, V. M. (2010). The impact of a classroom intervention on grade 10 students' argumentation skills, informal reasoning, and conceptual understanding of science. *Journal of Research in Science Teaching*, 47(8), 952–977. <https://doi.org/10.1002/tea.20358>
- Von Aufschnaiter, C., Erduran, S., Osborne, J., & Simon, S. (2008). Arguing to learn and learning to argue: Case studies of how students' argumentation relates to their scientific knowledge. *Journal of Research in Science Teaching*, 45(1), 101–131. <https://doi.org/10.1002/tea.20213>
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. In M. Cole, & Trans (Eds.), *V. John-Steiner, S. Scribner, & E. Soubberman*. Cambridge, MA: Harvard University Press.
- Wagner, J. (1997). The unavoidable intervention of educational research: A framework for reconsidering researcher-practitioner cooperation. *Educational Researcher*, 26(7), 13–22. <https://doi.org/10.3102/0013189X026007013>
- Walton, D. N. (1989). Dialogue theory for critical thinking. *Argumentation*, 3(2), 169–184. <https://doi.org/10.1007/BF00128147>
- Walton, D. (2011). Reasoning about knowledge using defeasible logic. *Argument & Computation*, 2(2–3), 131–155. <https://doi.org/10.1080/19462166.2011.637641>
- Weinberger, A., Stegmann, K., & Fischer, F. (2010). Learning to argue online: Scripted groups surpass individuals (unscripted groups do not). *Computers in Human Behavior*, 26(4), 506–515. <https://doi.org/10.1016/j.chb.2009.08.007>
- Wertsch, J. V., & Toma, C. (2012). *Discourse and learning in the classroom: A sociocultural approach*. In L. P. Steffe, & J. Gale (Eds.), *Constructivism in education* (pp. 177–192). London, UK: Routledge.
- Wilkinson, I. A. G., Reznitskaya, A., Bourdage, K., Oyler, J., Glina, M., Drewry, R., et al. (2017). Toward a more dialogic pedagogy: Changing teachers' beliefs and practices through professional development in language arts classrooms. *Language and Education*, 31(1, S1), 65–82. <https://doi.org/10.1080/09500782.2016.1230129>
- Windschitl, M., Thompson, J., & Braaten, M. (2008). Beyond the scientific method: Model-based inquiry as a new paradigm of preference for school science investigations. *Science Education*, 92(5), 941–967. <https://doi.org/10.1002/sce.20259>
- Wolfe, C. R. (2011). Argumentation across the curriculum. *Written Communication*, 28(2), 193–219. <https://doi.org/10.1177/0741088311399236>
- Zemal-Saul, C., Munford, D., Crawford, B., Friedrichsen, P., & Land, S. (2002). Scaffolding preservice science teachers' evidence-based arguments during an investigation of natural selection. *Research in Science Education*, 32(4), 437–463. <https://doi.org/10.1023/A:1022411822951>
- Zohar, A. (2008). Science teacher education and professional development in argumentation. In S. Erduran, & M.-P. Jimenez-Aleixandre (Eds.), *Argumentation in science education. Perspectives from classroom-based research* (pp. 245–268). Dordrecht, Netherlands: Springer.
- Zohar, A., & Nemet, F. (2002). Fostering students' knowledge and argumentation skills through dilemmas in human genetics. *Journal of Research in Science Teaching*, 39, 35–62. <https://doi.org/10.1002/tea.10008>