Educational Psychology Review (2022) 34:477–509
https://doi.org/10.1007/s10648-021-09637-2

REVIEW ARTICLE

Learning to Argue Through Dialogue: a Review of Instructional Approaches

Chrysi Rapanta[1] · Mark K. Felton[2]

Accepted: 12 August 2021 / Published online: 4 September 2021

Abstract
Over the past 20 years, a broad and diverse research literature has emerged to address how students learn to argue through dialogue in educational contexts. However, the variety of approaches used to study this phenomenon makes it challenging to find coherence in what may otherwise seem to be disparate fields of study. In this integrative review, we propose looking at how learning to argue (LTA) has been operationalized thus far in educational research, focusing on how different scholars have framed and fostered argumentative dialogue, assessed its gains, and applied it in different learning contexts. In total, 143 studies from the broad literature on educational dialogue and argumentation were analysed, including all educational levels (from primary to university). The following patterns for studying how dialogue fosters LTA emerged: whole-class 'low structure' framing with a goal of dialogue, small-group 'high structure' framing with varied argumentative goals, and studies with one-to-one dialectic framing with a goal of persuasive deliberation. The affordances and limitations of these different instructional approaches to LTA research and practice are discussed. We conclude with a discussion of complementarity of the approaches that emerged from our analysis in terms of the pedagogical methods and conditions that promote productive and/or constructive classroom interactions.

Keywords Argumentation · Dialogue · Review · Learning to argue · Patterns · Instructional approach

Over the past 20 years, a broad and diverse research literature has emerged to address how students learn to argue through dialogue in educational contexts. This field has sprung, in part, from a proliferation of research into the benefits of argumentation for learning (see Andriessen and Baker 2014; Asterhan and Schwarz 2016), which is based on the view that argumentation

✉ Chrysi Rapanta
crapanta@fcsh.unl.pt

[1] Faculty of Social Science and Humanities, Universidade Nova de Lisboa, Lisbon, Portugal

[2] Connie L. Lurie College of Education, San Jose State University, San Jose, CA, USA

supports knowledge construction, by situating learned facts, or knowledge, within evidence-based explanatory frameworks (Leitão 2000). In fact, extensive research has shown that by engaging in argumentative practices students from childhood to adulthood learn how to better structure their reasoning, consider alternative viewpoints (e.g., Kuhn et al. 2016a; Larrain et al. 2019; Nussbaum and Edwards 2011; Reznitskaya et al. 2009), and develop more nuanced understandings of the topics they argue about (e.g., Asterhan and Schwarz 2007; Iordanou et al. 2019; Zohar and Nemet 2002). Moreover, several scholars (e.g., Rapanta 2021; Kuhn 2018, 2019; Osborne et al. 2013) have confirmed that argumentation can be used as a pedagogical method to help students think critically, solve problems, and make decisions that concern their lives as individuals, as members of a society, and as world citizens.

Argumentation may involve an individual discourse activity, like argumentative writing (Klein et al. 2019) or graphical representation (Noroozi et al. 2018; Nussbaum 2008), or it may take the form of a social discourse activity in which dialogue partners collaboratively construct, critique, and reconcile arguments in the service of knowledge construction (Kuhn 2015). This view, rooted in social constructivist theories of learning, explains how skills of argument emerge and are refined on the social plane before becoming internalized to manifest as individual reasoning and learning (Billig 1987; Kuhn and Crowell 2011; Mercer and Howe 2012). However, for this socially rooted learning to take place, an immersion in dialogue is necessary (Cavagnetto 2010; Prawat 1991). According to Chinn and Clark (2013), when students learn how to argue through dialogue, they learn 'how to engage effectively in the practice of argumentation' (p. 321).

The current research is divided regarding what this 'effective engagement' entails and what it takes for researchers and educators to achieve it (Asterhan et al. 2020; Kim and Wilkinson 2019). On one extreme, there is the non-instrumental view, in which dialogue represents a form of social meaning-making, valuable in itself and not defined by external gains or pre-defined goals (Clarà 2021). From this perspective, learning to argue naturally emerges from adopting dialogic moves that facilitate the exchange of ideas or perspectives, such as asking open-ended questions, elaborating on others' ideas, exploring competing viewpoints, coordinating views, and reflecting on dialogue metacognitively (Hennessy et al. 2016; Howe et al. 2019). On the other extreme, there is the intentional instructional framing of dialogue to produce specific argumentative gains. From this perspective, learning to argue involves using dialogue to define problems, produce claims, critiques and rebuttals using evidence, and coalesce arguments in ways that produce more accurate, developed, or sound outcomes (Baker 1999; Chen et al. 2016; Zohar and Nemet 2002).

Adopting the non-instrumental point of view, engaging in dialogue is the primary focus on LTA, whereas adopting the argument-oriented point of view, engaging in argumentation is the primary focus of LTA. Independently of the approach taken, dialogue needs to have certain characteristics and/or lead to certain gains so that it can fulfil its learning potential (Resnick et al. 2018). Although extensive research is available on the educational dialogue that leads, directly or indirectly, to LTA outcomes, these outcomes are not defined similarly across studies, and the methods adopted to define and measure them do not always involve controlled comparison to traditional teaching methods, so that valid conclusions about the effectiveness of instructional approaches can be drawn (Hoadley 2006). Our goal with this review is to better distinguish and define the range of approaches taken by researchers and teachers to elicit and advance students' reasoning in the social context of classroom interaction, with effectiveness being defined either in terms of interventions *on* the quality of discourse and dialogue, or as a characteristic *of* the dialogue itself. By characterizing and mapping these various approaches to

LTA, we hope to contribute to the current discussion of how to promote students' argumentation competence through interactive discussions (Noroozi et al. 2018). In the section that follows, we provide definitions of dialogue and argumentation, as well as our own conceptualization of LTA from an instructional framing perspective.

Defining Dialogue and Argumentation

From a social constructivist perspective, dialogue is a critical activity through which knowledge construction, in the broad sense of meaning making, takes place (Ford 2012). Dialogue partners must commit to taking turns articulating their own ideas as speakers and seeking to understand their partner's ideas as listeners (Wells 2007). This *dialogicality* (Koschmann 1999) introduces a tension between the effort to enter into a shared understanding (*intersubjectivity*) and the effort to distinguish one's own thoughts from another's (*alterity*). As a result, knowledge is constructed between speakers, rather than being transmitted from one to the other. This basic idea of dialogic exchange cannot be taken for granted in a classroom context. When it comes to teacher-student interactions, research has shown that much of teacher discourse is monological (Scott et al. 2006), meaning the dynamic and equitable exchange of roles between speakers does not take place. Similarly, when it comes to student-student interaction, much of the classroom dialogue is either disputative or cumulative (Mercer 2004), neither of which leads to the productive resolution or contrast of perspectives.

Argumentative dialogue is a dialogue activity in which the tension between intersubjectivity and alterity is addressed by the balance, constantly constructed by participants, between construction and critique (Ford 2008). On one hand, dialogue participants use language to construct and support a contested or potentially contested assertion (Rapanta 2019; Rapanta and Macagno 2019); on the other hand, they continuously strive to coordinate between the evidence supporting their own position and evidence supporting alternative or contrary positions (Kuhn 1999). In other words, argumentative dialogue involves an exchange 'in which the participants not only defend their own claims, but also engage constructively with the argumentation of their peers' (Nielsen 2013; p. 373). In so doing, they propose and critique the grounds for accepting a possible resolution to the doubt or uncertainty driving argumentation. Productive engagement, aimed at the resolution of doubt or uncertainty, creates the potential for argumentative dialogue to lead to knowledge construction, as speakers engage deeply with each other's thinking in a critical discussion (Keefer et al. 2000). However, there is always the risk of a 'position-driven' argument, where authentic efforts at reasoning about opposing claims and evidence are replaced by strategic efforts at winning the exchange regardless of the quality of the arguments under consideration (Keefer et al. 2000).

It is because of this dialectical nature—which, as we will show further on, can be more or less explicit depending on how dialogue is framed—that argumentative dialogue can have more possibilities than any other type of dialogue to be both constructive and critical. This quality, as a result, brings several benefits at a social, cognitive, and epistemological level, briefly described below.

Social Outcomes

Research that emphasizes social gains as a result of engagement in argumentative dialogue operates on the assumption that a so-called 'argument-as-process' emerges from dialectical

exchanges between speakers who hold opposing views on a topic. From this perspective, divergent views fuel a form of discourse in which individuals draw out and challenge the claims and evidence for each other's conclusions (Felton and Kuhn 2001; Kuhn and Zillmer 2015). Focusing speakers on the differences in their views encourages them to explore the relative strengths of arguments in ways that are often absent when speakers focus on similarities in their perspectives (Thiebach et al. 2016). As a result, the social dynamic of dialectical exchange creates a context in which speakers naturally prompt one another to produce more fully elaborated and carefully examined arguments than might be later observed in independent reasoning (Kuhn et al. 1997; Kuhn 2019; Larrain et al. 2019). This results in the development of the social, critical thinking skill of 'antilogos', that is, one's ability to identify limitations in one's own assumptions and positions, which may lead to totally different or even oppositional assumptions and positions (Billig 1987).

Cognitive Outcomes

Several higher-order cognitive and metacognitive processes are inherent in argumentative dialogue due to the complexity and diversity of the task of putting forward a position, defending it with sufficient and relevant evidence, addressing counterarguments, generating valid rebuttals to further strengthen one's own side, or making revisions to one's argument in light of valid critique. Together, these activities draw on a set of cognitive, meta-cognitive and meta-strategic resources (Rapanta et al. 2013), including inferential abilities, evaluation abilities, argument appraisal, as well as more advanced strategies of theory-evidence coordination and undermining an opponent's point of view with counterevidence. In a climate of fake news and alternative facts, learning to identify, assess, and integrate information from sources to draw valid conclusions is a direct benefit of argumentative dialogue.

Epistemological Outcomes

Argumentative reasoning, including both the construction and evaluation of arguments, and epistemic cognition are highly intertwined (Chinn et al. 2011; Iordanou et al. 2016). Argumentative dialogue, as a social manifestation of argumentative reasoning, helps epistemic cognition to become explicit, especially when it comes to the acquisition of norms of what counts as a good argument and argumentative dialogue (Kuhn et al. 2013). As argumentative dialogue externalizes thinking, it creates a context for reflective judgments about the validity of each speaker's claims and evidence. As a result, it opens the door to new insights into the nature of knowing, standards for judging certainty, and practices for constructing knowledge from carefully vetting claims and evidence.

Although not all educational dialogues are aimed at argumentation as a goal, argument-related gains like the ones described above are a part of the learning potential of dialogue (Resnick et al. 2018). However, when dialogue is not designed to be argumentative, those social, cognitive, and epistemological gains often emerge as desired by-products or even manifestations of the dialogic process, even though they were not the direct object/-ive of dialogue. To be able to bridge the two approaches, i.e., the more instrumentalist one, directly aiming at argument gains, and the more ontological one, which welcomes argument gains as part of the dialogic engagement (Clarà 2021), we provide below a conceptualization of LTA instructional approaches which applies to both cases.

LTA Instructional Approaches

According to Morrison et al. (2019), the fundamental components of any instructional design, no matter what the designed learning activity/environment is, are: the Objectives, which 'provide a map for designing the instruction and for developing the means to assess learner performance' (p. 17); the Methods, which define how the subject content and/or skills are best learnt; the Learners, referring to the characteristics of learning for whom the activity/environment is designed; and the Evaluation, referring to ways to determine the extent to which learning is achieved. Applying this instructional effectiveness model directly to the LTA literature is not possible, due to the great variety of approaches that exist within as discussed above. Therefore, what we propose is to transform the generic instructional effectiveness model by Morrison et al. (2019) into a framework for reviewing the LTA literature, which can adapt to the diversity of approaches encountered within. Our MeDOL framework adapts Morrison et al.'s (2019) effectiveness components as follows: Methods refer to the authors' approach of framing educationally effective dialogue, and to the fostering methods used to increase LTA outcomes; Dialogue goals refer to the epistemic goals pursued by the study participants during the LTA activity; Outcomes refer to the participants' LTA gains described in each study; and Learners refer to the students and learning main characteristics described in each study such as age/educational grade, learning structure (i.e., whole class, small groups, dyads, or mixed), and disciplinary field.

Methods

It has been argued that framing methods shape, or at least characterize, the process by which educational goals are achieved through dialogue (Ford and Wargo 2012). When the goal is LTA, dialogue can either be framed a priori as argumentation, or emerge as a result of the dialogic activity. In this sense, argumentative dialogue may emerge through a more structured approach, where the learning environment is intentionally designed to produce specific argumentative interactions and outcomes; or it may follow a less structured approach, where the learning environment is designed to promote dialogic norms, without specifically focusing on the production of particular argumentative interactions or outcomes, as the ones described above. We will henceforth refer to this distinction as 'high vs low structure' approaches to argumentative dialogue.

Low-structure approaches to educational dialogue do not explicitly focus on structuring argumentation as a goal-oriented activity itself, but rather on establishing norms for dialogue and its productivity. In this approach, educational (argumentative) dialogue is productive insofar as participants actively engage in exploring different perspectives. Argumentation may emerge as a natural result of exploring these perspectives, but it is not set as the primary outcome. An example of this approach is 'exploratory talk' (Mercer 2004).

In contrast to low-structure dialogic approaches, the high-structure ones focus on argumentation as an explicit goal, and not a by-product of students' interactions. This is primarily done through focusing on argumentative knowledge construction, therefore on the types of dialogic activities students should engage in so that some type of new knowledge or understanding can emerge. This type of educationally effective dialogue is particularly common within the fields of computer-supported collaborative learning (CSCL) (see, for example, the argument scripting approaches—Jermann and Dillenbourg 2003) and science education (see, for example, the Science Writing Heuristic approach—Hand et al. 2021).

A separate case of high-structure dialogic approaches focuses on the resolution of a controversy as the vehicle for promoting advances in argumentation. Within this approach, which can be called dialectic (Asterhan 2013), the characteristics of the issue and the nature of the task assigned to students as part of its resolution are of high priority in terms of designing the interaction. Regarding the issue, the more controversial it is, the greater the possibilities that authentic argumentation will emerge, in the form of a critical discussion (van Eemeren and Grootendorst 1992; Walton and Krabbe 1995). Moreover, a clear need for the issue resolution must be present either in the form of a dilemma (e.g., Kuhn 2018; Zohar and Nemet 2002), or in the form of a decision among multiple alternatives (e.g., Jiménez-Aleixandre 2002; Garcia-Mila et al. 2013).

Dialogue Goals

From a dialogue theory perspective, the identification of types of dialogue according to dialogue goals has attracted researchers' attention (see, for example, Walton 1989, 1998). When it comes to educational contexts, it is still not clear what such a typology should take into consideration as a 'goal'; therefore, different 'goal' approaches result into different dialogue types. For example, Osborne et al. (2016) proposed two main interlinked goals for science argumentation: constructing and critiquing scientific explanations. Keefer et al. (2000) propose four types of educational dialogues by crossing 'convergent' or 'divergent' participant approaches, on one hand, with how successful they are in fulfilling their goal, giving: (a) a successful divergent approach called 'critical discussion' (term borrowed from Van Eemeren and Grootendorst 2003), (b) a non-successful divergent approach called 'eristic discussion', (c) a successful convergent approach called 'explanatory inquiry', and (d) a non-successful convergent approach called 'rapidly reaching consensus'.

In our view, both proposals discussed above are insufficient to capture the complexity of dialogue goals in LTA instructional settings. On one hand, Osborne et al.'s (2016) proposal clearly focuses on the epistemic aspects of instructional framing for LTA without refining the dynamics of knowledge construction and critique. On the other hand, Keefer et al.'s (2000) focuses on the dialogue fluidity and versatility 'on the go', and not as part of an instructional framing. We view dialogue goals as epistemic goals enacted by the student participants in each dialogic context. Examples of such dialogue goals are the ones proposed by Berland and Reiser (2009), namely: sensemaking, articulation, and persuasion. Extending Berland and Reiser's (2009) initial proposal, *sensemaking* dialogues aim at making sense of what others say and/or a specific phenomenon under consideration. *Articulation* focuses on either articulating the relationship between sources of evidence or coordinating theories with evidence. It is about understanding at least two types of data and establishing the relationship between them. Finally, persuasion dialogues focus on either arriving at a consensus (deliberation) or proving one's side or position as the best according to epistemic standards. When it comes to persuasion as an ultimate goal in educational dialogue, a deliberation phase is always necessary, particularly when a final decision must be made. For this reason, and also to avoid the misconception that persuasion must be confrontational (Kruger 1993; Micheli 2012), we can refer to this goal as a deliberation goal, rather than persuasion

Outcomes

The traditional distinction between *argument1* and *argument2* (O'Keefe 1992) proposes two legitimate and complementary approaches to argument, one that focuses on

argument-as-product (i.e., something that a person makes) and another that focuses on *argument-as-process* (i.e., something that a person engages in).

Arguments-as-products may comprise both structural and functional elements of discourse (for more about this distinction, see Rapanta et al. 2013; Macagno 2016). In terms of structural elements, effective argumentative reasoning might take some of the following forms: (a) effective integration and use of evidence (i.e., Berland and Reiser 2011; Kuhn et al. 2013); (b) effective integration of arguments and counterarguments, i.e., balanced or dialogical arguments (i.e., Kuhn and Udell 2007; Polo et al. 2016); and (c) effective use of counter-arguments and rebuttals (i.e., Kuhn and Udell 2003; Crowell and Kuhn 2014). It may also relate to the elaboration of arguments, as with the structural elements proposed by Toulmin (1958). More precisely, *data, warrants*, and *backings* are related to the identification and use of evidence, whereas *claims, qualifiers*, and *rebuttals* are related to the skill of construction and critique. Of course, to be able to assert that the use of specific argument elements is effective, other pragmatic criteria may be needed, such as conceptual complexity, coherence, and relevance.

In terms of argumentative function, effective dialogue takes the form of discursive moves that elicit reasoning, operate on reasoning, or redirect conversation. Early work in the transactive nature of dialogic reasoning (Berkowitz and Gibbs 1983; Kruger 1993) plays an important role in many functional models of argumentative reasoning. This work focuses on the ways in which speakers engage with and operate on each other's thinking through dialogue, with an emphasis on how dialogue can be used as a way to socially elicit, elaborate, critique, and revise arguments. Generally, analyses of discursive moves in argumentation focus on the ways in which instructional interventions, scaffolds, or contextual variables facilitate these discursive processes. Some studies also go beyond the analysis of individual moves to look at how these moves are coordinated to form either coherent dialogic processes or effective argumentative strategies. It is common to speak of 'dialogic transactivity' in the first case, and of 'dialectic transactivity' in the latter (Vogel et al. 2016).

In recent years, social scientists have begun to challenge the notion that good argumentation involves distilling rational argument from the soup of social and emotional conflict. Instead, argumentation is seen as a dynamic interplay of social, emotional, and cognitive functions (Asterhan 2013; Gilbert 2004; Plantin 2004). On the one hand, emotions play an important role in promoting active engagement in reasoning and complement rationality (Lipman 2003). On the other hand, there is clear evidence that negative emotions drive the distortion of facts, increase miscommunication, and lead arguments astray in disputative argumentation (Polo et al. 2016). For example, Polo and her colleagues (Polo et al. 2016) propose that attempts to maintain positive emotion may be the reason that cumulative talk fails to draw out the kind of constructive criticism found in exploratory talk. Therefore, the ability to modulate and navigate emotions in a dialogue is an important component of argumentation (Andriessen et al. 2011).

There is also a social dimension to the regulation of argumentation, since the norms of social interaction can also have a direct impact on the progress of argumentative dialogue. In fact, there seems to be a close relationship between norms for social interaction and emotional tension in argumentative dialogue. At the most basic level, turn-taking is a social norm that, when broken, can negatively impact dialogue. When speakers talk over one another, cut each other off, or dominate the conversation, they undermine the appearance of good faith, and threaten other speakers' face, i.e., perceived public image (Chiu 2008). Conversely, positive social regulation, such as active listening among peers, and between teacher and students (Alexander 2017; Michaels et al. 2008), can serve to facilitate effective argumentative dialogue by creating a safe space for the critical evaluation of ideas.

Finally, metadialogue (Krabbe 2003) is a socio-epistemic regulator of argument-as-process that contributes to effective argumentation and at the same time functions as a manifestation of argument gains. It may include appeals to standards for valid reasoning, where speakers make an explicit statement about the nature of claims and evidence and the ways in which they are coordinated to draw a conclusion (Macagno et al. 2015). Metadialogue can also be used to regulate the goals, process, and outcomes of argumentative dialogue (Felton et al. 2015a, b): speakers may explicitly coordinate goals, focus dialogue on points of disagreement, or propose solutions to apparent contradictions in their views. Finally, metadialogue can be used to impose social norms for regulating argumentative dialogue (Kuhn and Zillmer 2015; Michaels et al. 2002). Across these disparate applications, metadialogue can be understood as the conscious and explicit attempt to optimize the social construction of arguments. Seen in this way, metadialogue represents a form of epistemic cognition operating on the process of argumentation (Kuhn et al. 2013). Indeed, metacognitive knowledge and regulation of argumentation are naturally epistemological because of the role that argumentation plays in establishing the strength, validity, truth, or applicability of conclusions through the coordination of claims and evidence. When speakers choose to direct or correct the course of argumentation with metadialogue, they manifest epistemological aims, knowledge, and values about argument.

Learners

Finally, all the above (Methods, Goals, Outcomes) makes sense because it works for a concrete type of learner in a concrete context. The learner characteristics expected to be found in all empirical reviewed studies are: (a) the age/educational grade; (b) the disciplinary field or topic on which the dialogue is held; and (c) the setting of the learning situation, meaning whether the dialogue take place in whole-class discussion format, in small groups, or in one-to-one settings.

The Present Study

To address the existing tension between 'dialogue' as the main focus and 'argumentation' as the main focus of research, we propose to look at LTA as a common instructional process to pursue either explicitly or as a by-product. Our approach is different to theoretical approaches focusing on dialogue purposefulness (Alexander 2017) or argumentation goal-orientedness (Walton 2013): rather than looking at the concrete aims participants strive at during their engagement in argumentative dialogue, we opt for studying the different ways in which learners' engagement in the practice of argumentation is framed, fostered, and evaluated by educational researchers. We propose that by studying the literature in this way, we might better understand how different approaches to LTA might inform and complement one another. We believe that doing so might bring more coherence to our understanding of how to foster argumentative dialogue and to what ends.

Our research questions are:

1. How do educational researchers frame argumentative dialogue and its gains in instructional settings?
2. How is argumentative dialogue scaffolded and fostered in different contexts?

Methods

The integrative review was selected as the most appropriate way to address the research problem of mapping existing educational research focusing on learning to argue (LTA) through dialogue. Given the complexity of argumentative discourse and the breadth of approaches taken to describe and study the phenomenon, we decided to include both experimental and non-experimental studies in our review, whereas a meta-analysis would limit our scope to experimental studies only (Whittemore and Knafl 2005). In addition, our goal was not to measure the effectiveness of dialogue on students' learning (such a goal would have been impossible given the variety of methodological approaches), but to understand how learning to argue is actually carried out in different instructional contexts. We therefore aimed at studies from both educational dialogue and argumentation fields, while limiting our search to those studies that offered an evaluation of the quality of discussions revealing or leading to some type of learning to argue gains. The indicators we used for the LTA evaluation are included in the search keywords described below.

Sample

A systematic search of peer-reviewed articles was conducted using four large databases: ISI's Web of Science (WOS), Elsevier's Scopus, EBSCO, and PROQUEST. We intentionally aimed at studies explicitly focusing on some manifestation and assessment of dialogue quality, defined in a variety of ways, such as productive, constructive, effective, strategic, or persuasive. For each one of the searches, the keywords used were: 'argument*' or 'dialog*', on the one hand, and 'effect*' or 'construct*' or 'productive' or 'strateg*' or 'persuas*' as LTA quality indicators, on the other.

The initial search yielded 3013 results, excluding unpublished manuscripts, book chapters, and conference proceedings, to ensure a higher quality of the studies due to stricter peer review criteria. A first screening of these documents based only on the title resulted in a sample of 945. This reduction was mainly due to the frequent use of the term 'argument' in its ordinary, non-technical sense. The pool was further reduced after the exclusion of duplicates, as several databases were simultaneously considered, to 761 articles. A second screening based on the abstract resulted in 341 articles. For this second screening, we used the following inclusion criteria: (a) that the article was empirical rather than theoretical or methodological; (b) that the article addressed K-16 education; (c) that the study addressed teacher-students and/or student-student interactions in either face-to-face or computer-supported classroom contexts. The selected articles underwent a third screening based on their full text. To the previous inclusion criteria, the following two were added: (a) that the quality of dialogue and/or reasoning were assessed in some way; and (b) that the interactions were held in the classroom or other educational setting (e.g., computer laboratory). This final screening resulted in 145 articles, with two cases of pairs of articles reporting on the same studies; therefore, we ended with a total number of 143 studies published between 1997 (date of the first included source) to 2020 (until November, and only finally published documents).

Coding Process

All 143 studies were coded in terms of the MeDOL framework presented in the theoretical part of this paper. Eleven coding categories emerged from our analysis (Figure 1). Four of these

Coding categories	**Sub-categories** (when applicable; asterisk '*' indicates categories openly coded)
Methods used to frame and foster dialogue	
Dialogue approach	Choose **one** of the following that best describes the dialogue approach taken by the authors: (a) low-structure dialogic; (b) high-structure dialogic; and (c) dialectic.
Coach-based*	Describe any concrete move, stance or technique (e.g. modelling discourse) used by the instructor to explicitly foster argumentative dialogue.
Task-based	Indicate which one(s) of the following task components have been shown as relevant or significant for the quality of argumentative dialogue: (a) the issue(s)/topic discussed; (b) the class organisation (seating/group structure); (c) how the task is organised in terms of contents and actions (task structure/script); or (d) a specific instructional technique adopted as a task activity (e.g. philosophical circle, Science Writing Heuristic approach, etc.).
Materials-based*	Indicate whether there was a specific material resource (visual, online, other) that was considered relevant or significant for fostering argumentative dialogue
Peers-based*	Indicate any peers' move or role (enacted or pre-assigned) that was considered to foster the argumentative dialogue quality.
Dialogue goal	Choose **one** of the following that best describes the epistemic goal set for the argumentative dialogue: (a) sensemaking (explanation); (b) articulation (theory-evidence coordination); (c) deliberation (persuasive deliberation)
Outcomes	
Argument-as-product	Choose **one** of the following: (a) main argument, when reasoning quality is assessed based on claims and/or reasons and/or the type of claims; (b) elaborated, when the focus is on use of evidence and/or Toulmin's (1958) argument structure; (c) complex, when the focus is on the internal relationships of argument elements and/or quality (types) of evidence/justification; (d) dialogic, when the focus is on transactivity and/or use of critique and rebuttal and/or use of justification for critique/rebuttal; and (e) dialectic, when the focus is on revision or integration of alternative theories within one's own argumentation and/or types of counterarguments/rebuttals and/or use of critical questions.
Argument-as-process*	Indicate any type of: (a) socio-emotional factors or processes that function as indicators of argumentative dialogue quality; and (b) meta-level (cognitive, thinking, discursive) processes or dispositions that function as indicators of argumentative dialogue quality.
Learners	
Dialogue setting	Choose **one** of the following: (a) one-to-one setting (where 'one' corresponds to a unit or a pair); (b) small groups; (c) whole class; (d) mixed.
Educational grade	Choose **one** of the following: (a) primary (approx. 5-11 years old); (b) secondary (approx. 12-17 years old); (c) University (above 17); (d) combined (more than one age levels)
Subject field	Choose **one** of the following: (a) science; (b) socio-scientific issue; (c) mathematics; (d) language; (e) history; (f) general; (g) combined/other

Fig. 1 The coding scheme (Note: Categories marked with an asterisk were open-coded)

variables (coach-based, materials-based, peers-based fostering methods, and argument-as-process) were openly coded, and the remaining seven (dialogue approach, task-based scaffolds, dialogue goal, argument-as-product, dialogue setting, educational grade, and subject field) were defined a priori, with sub-categories explained below.

During the coding process, the two authors randomly selected 20% of the studies and double-coded them independently for three high-inference variables, namely dialogue approach (low structure/high structure/dialectic), dialogue goal (sensemaking/articulation/deliberation), and argument-as-product (main/elaborated/complex/dialogic/dialectic). The inter-rater reliabilities were satisfying for all three variables (κ = .776, κ = .727, and κ = .702, respectively), and disagreements were resolved by discussion.

Dialogue Approach

As mentioned in the theoretical part of this review, three sub-categories were used to describe the different approaches to LTA: (a) low-structure dialogic, (b) high-structure dialogic, and (c) dialectic. A main distinction between the *high* versus *low structure* approaches to argumentative dialogue is that in the low structure approach, dialogic interaction has the predominant 'formative potential' (Chin and Teou 2009), shaping students' ideas and thinking by creating a shared dialogic ethos (Littleton and Mercer 2013). In contrast, in the high structure approach, it is the design of the learning environment that supports students' enactment of meaningful argumentative practices (Ravenscroft 2000; Wu and Krajcik 2006). Similarly, a main distinction between the dialectic and the high structure dialogic approach is that the high structure guides students through a defined set of procedures to reach a specific outcome (e.g., an answer to a problem), whereas the dialectic approach focuses on eliciting opposing views to encourage each party to consider their arguments in a framework of alternatives.

Task-based Scaffolds

These refer to task components that have been shown to be relevant or significant for the quality of argumentative dialogue in the various studies. Among the task-based scaffolds, we pre-defined the following sub-categories: (a) the issue(s)/topic discussed; (b) the class organization (seating/group structure); (c) how the task is organized in terms of contents and actions (task structure/script); or (d) a specific instructional technique adopted as a task activity (e.g., philosophical circle, Collaborative Reasoning, etc.).

Dialogue Goal

As mentioned in the theoretical part of this review, three sub-categories were used to describe the different dialogue goals of LTA: (a) sensemaking, (b) articulation, and (c) deliberation/deliberative persuasion. According to Dougherty et al. (2000), 'people cannot collectively use knowledge unless they first make shared sense of it' (p. 323). Therefore, *sensemaking* is a starting point for any constructive dialogue (i.e., a dialogue that explicitly aims at constructing new knowledge) to take place. When learners engage in 'sensemaking' as their main task, they question each other, clarify content, build on each other's ideas, and share their individual explanations about a phenomenon. As part of this task-process, several cognitive processes can take place such as: describing multiple aspects of a phenomenon, comparing two or more phenomena, or describing different types/cases of a phenomenon (Meyer et al. 2015).

Examples of studies focusing on sensemaking as a primary task for students to engage with are: Lee, Kang, and Kim (2015)[1] in science (students were asked to provide an explanation of a scientific phenomenon using modelling strategies) or Chisholm and Loretto (2016) in a language classroom ('how students made meanings in interaction by dialoguing with other students, texts, and ideas', p. 1).

[1] It is interesting to note that Lee et al. (2015) used the triadic model (sensemaking-articulation-persuasion) proposed by Berland and Reiser (2009) as a method of analysing goals emerging in students' discourse. This view of the three tasks as individual cognitive operations is different than the approach adopted in this study, where we view the three tasks as social activities held dialogically in the classroom.

Articulation focuses on either articulating the relationship between sources of evidence and/or coordinating theories with evidence. It is about understanding at least two different types of data and establishing a relationship between them. In articulation-oriented dialogue tasks, students exchange, contrast, and coordinate their interpretations of a phenomenon using some type of data (information) as evidence. Based on Baker (1999), Veerman et al. (2000) define articulation as a knowledge transformation process, during which 'already stated information is evaluated and integrated into the collective knowledge base in such a way that, a new insight or a new direction transpires, that can be used to answer questions or to solve problems' (p. 272). Although sensemaking prepares the ground for joint knowledge construction and learning with others, articulation makes this goal explicit. Examples of studies focusing on articulation as a primary task for students to engage with are: Kim and Song (2006) in science (all students were asked to defend their scientific evidence-based reports in front of the rest of their classmates who would challenge it with questions, without the goal being that of choosing the report that best explained the phenomenon at hand), or Jadallah et al. (2011) in a language classroom ('One of Ms. Jackson's principal objectives was to have children support their arguments with story evidence', p. 204).

Finally, *deliberation* focuses on either arriving at a defending one's position on a topic or reaching consensus after weighing alternatives based on epistemic standards. The activities characterizing deliberation dialogue involve addressing explicit disagreement to reach a conclusion or decision or final state of a debate, while revising one's position as necessary. Defining deliberation as the main task of learning to argue implies an explicit focus on increasing the plausibility of one position, either through accepting it as the most accountable one (e.g., by consensus or compromise) or through persuading the other party using logical arguments. Sensemaking and articulation are both elements of the deliberative process, as opposing claims must be first elaborated and then substantiated in order to be critically evaluated. Examples of studies focusing on persuasion/deliberation as a primary task for students include Felton et al. (2015a, b), where science students used critical dialogue to evaluate solutions to a socio-scientific problem, or Muller-Mirza et al. (2007)[2], where students engaged in a historical debate through role-playing.

Argument-as-Product

The sub-categories and codes we used to capture the argumentative reasoning outcomes, when these were expressed as functional or structural products of discussion (see theoretical part of this review) were: claim (C), reason (R), claim quality (Cq), evidence (E), questions (Q), Toulmin's (1958) Argument Pattern (TAP) structure, e.g., claim-data-(warrant)-backing, connection between claim and reasons (C/R), evidence quality/types (Eq), types or quality of reasons (Rq), transactive moves (T), counterargument and/or rebuttal (CAR), evidence-based counterargument and/or rebuttal (CAR/E), revised or integrated argument (RE), quality/types of counter-argumentation strategies (CARq), use of critical questions by students (Qq). Table 1 shows the description of the five levels of argument quality, described previously, based on the presence or combination of the above elements.

[2] It is worth noting here that this study, and others in our sample, did not identify learning to argue as their main pedagogical goal. Muller-Mirza et al. (2007) explicitly claim that their goal 'was not to teach students how to argue, but to provide them with the opportunity to learn from argumentation' (p. 256). However, the study was still included because there was enough information provided regarding how students engaged in the practice of argumentation and the gains achieved from this engagement, other than content learning.

Table 1 Argument quality levels used for the coding of argument reasoning gains category

Level	Codes	Argument quality type
1	C, R, Cq	main argument: focus on presence of claims and/or reasons and/or the type of claims
2	E, Q, TAP	elaborated argument: focus on use of evidence, questions, and/or TAP structure
3	C/R, Eq, Rq	complex argument: focus on the internal relationships of argument elements and/or quality (types) of evidence/justification
4	T, CAR, CAR/E	dialogic argument: focus on transactivity and/or use of critique and rebuttal and/or use of justification for critique/rebuttal
5	RE, CARq, Qq	dialectic argument: focus on revision or integration of alternative theories within one's own argumentation and/or types of counterarguments/rebuttals and/or use of critical questions

Findings

Study Demographics

In terms of the research design employed, the 143 studies comprised 82 (57%) descriptive, 37 (26%) quasi-experimental, and 24 (17%) experimental studies. When it comes to students' educational grade, 43 out of the 143 studies (30%) focused on primary school children, 43 (30%) on middle grades, 26 (18%) on secondary, 28 (20%) on university students, and a remaining 3 (2%) on more than one grade. Finally, regarding disciplinary field or type of issue discussed by students, 61 studies (43%) were about science, 32 (22%) were about general interest/social issues, 17 (12%) were about language/literature, 11 (8%) were about mathematics, 11 (8%) were about socio-scientific issues, 10 (7%) were about a combination of fields, and 1 was about history.

Instructional Framing Approaches to LTA

To explore patterns in framing argumentative dialogue, we looked at the relationship between 'dialogue approach' and 'dialogue task,' testing their independence using Fisher's exact test, with Cramer's V (φc) to look at effect size. Adjusted standardized residuals of +/−2 were used to determine the contribution of individual cells in the omnibus tests (Beasley and Schumacker 1995). The test yielded significant results ($p < .0001$), with a moderate effect size ($\varphi_c = .429$). Upon analysis of standardized residuals, the four relationships driving our findings were low-structured sensemaking, high-structured articulation, high-structured deliberation, and dialectic deliberation (Table 2). Due to the significant frequency of their cross-tabbed observation, these four instructional framing stories emerging from the analysed studies were identified as study patterns (see Sandelowski and Barroso 2007, for a detailed view of how patterns are defined in literature reviews).

Argumentative dialogue evaluation was approached in terms of gains in argumentative reasoning (argument-as-product) and/or in socio-emotional and metadialogic gains in argumentation (argument-as-process). In terms of reasoning gains, we had previously identified five types or levels of arguments, comprising both structural and functional elements, namely: main, elaborated, complex, dialogic, and dialectic (see Table 1). A significant association was found between the 'study pattern' and the 'reasoning gains' using Fisher's exact test, $p < .002$,

with a moderate effect size, $\varphi c = .329$. Table 3 shows the reasoning gains types reported per study pattern.

As can be observed in Table 3, a large majority of Pattern 2 (high-structured articulation) studies refer an elaborated structure of student arguments (i.e., a TAP, evidence-based claim structure), and a large majority of Pattern 4 (dialectic deliberation) studies focus on integrated arguments (i.e., integrating the dialectical argument-counterargument relationship), whereas a no clear association emerged for Pattern 3 (high-structured deliberation) and Pattern 1 (low-structured sensemaking) studies.

When it comes to the reported socio-emotional and metalevel gains, 48 out of the 108 pattern studies mentioned at least one of the two argument-as-process gains (i.e., socio-emotional and/or metalevel) as relevant for learning to argue. Of these, 17 studies exclusively focused on socio-emotional argument-as-process gains, 21 studies exclusively focused on metalevel gains, while eleven studies simultaneously considered both types of gains. These studies were distributed among Pattern 2 and Pattern 3 studies.

What types of different stories do the four patterns tell us regarding how learning to argue is achieved? Below we will give a descriptive account for each one of the four identified patterns, composed of different instructional approaches, tasks, and outcomes related to the 'umbrella' learning to argue (LTA) instructional goal.

LTA Pattern 1: Low-structured Sensemaking

Studies belonging to the first pattern (low-structured sensemaking) tend to describe argumentative dialogue as an organic, whole-class, student-driven (Aukerman et al. 2016) or student-dominant (McNeill and Pimentel 2010) dialogue, aiming at the co-construction of ideas and shared interpretive authority among teacher and students alike (Chisholm and Loretto 2016). Within this study pattern, argumentative dialogue is a learning conversation (Simon et al. 2008), in which student agency and authority *in* discourse (Forman et al. 2017) is promoted. These studies tend to define argument quality as a description of different modes of participation such as public warrantability (Atwood et al. 2010), 'interthinking' (Aukerman et al. 2016), agreeing/disagreeing (Topping and Trickey 2007), or divergent thinking (Damico and Rosaen 2009). When they focus on the structural elements of the arguments produced by students, this is mostly the warrant or explanation of reasoning (e.g., Coker Jr and Erwin 2011; Frijters et al. 2008; Langer-Osuna 2015; Langer-Osuna and Avalos 2015; Lee and Majors 2003). In very few cases, they present some socio-emotional or metalevel gains as result of students' participation in dialogue. Some of these gains are explained below.

Table 2 Cross-tabulation of dialogue task × dialogue approach[a]

		Dialogue approach			
		Low structured	High structured	Dialectic	Total
Dialogue task	Sensemaking	19[b]	12	2[b]	33
	Articulation	12	32[b]	2[b]	47
	Deliberation	7[b.]	25[b]	32[b]	63
	Total	38	69	36	143

[a] *0 cells (0.0%) have expected count less than 5*

[b] *Adjusted residuals >2 or <−2.*

Table 3 Reported reasoning gains of argumentative dialogue per study pattern ($N = 100^a$)

	Reasoning gains					
	Main	Elaborated	Complex	Dialogic	Dialectic	Total
Pattern 1 (low-structured sensemaking)	4[b]	3	3	3	3	16
Pattern 2 (high-structured articulation)	3	13[b]	7	5	4[b]	32
Pattern 3 (high-structured deliberation)	2	6	4	8	2[b]	22
Pattern 4 (dialectic deliberation)	1	2	3	7	17[b]	30
Total	10	24	17	23	26	100

[a] *Of the N = 108 studies belonging to one of the four dialogue study patterns, eight did not mention any argument reasoning gains, as they placed focus on other types of arguing skills (e.g., at a socio-emotional or meta-cognitive level)*

[b] *Adjusted residuals >2 or <−2*

Reznitskaya et al. (2012) describe a study in which Philosophy for Children (P4C) sessions provide a discussion frame suitable for the construction of new meanings and the reflection on the reasoning processes. Making sense of the concepts discussed and of the dialogue itself is therefore the primary goal set for the discussion. In the science context, Simon and her colleagues (Simon et al. 2008) describe how students' engagement in scientific talk during whole-class discussions can be fulfilled in ways that children talk about their ideas, clarify their thinking, and, consequently, develop their capacity to reason. In both examples, the gains of dialogue are not separate from the dialogue itself, adopting a participation rather than an acquisition approach to learning (Sfard 1998). In particular, Reznitskaya et al. (2012) characterize LTA dialogues as inquiry dialogues, placing the focus on the fact that there are no right and wrong answers during P4C discussions. Simon et al. (2008) largely define LTA dialogues as argumentative interactions, as opposed to recall responses, characterizing the former as learning conversations. An emphasis on low-structure instructional framing is evident in both definitions. Other studies within this pattern group highlight some manifestations of LTA gains, either as a product or as a process. For example, Topping and Trickey (2007) emphasize that students explicitly agree or disagree based on reasons, and for that to be possible, a process of increased student participation and responsiveness is recommended. Sutherland (2006) presents more advanced reasoning gains as manifestations of LTA dialogue, such as posing higher-order questions and producing critical responses. Veerman et al. (2000) found out that checking conceptual information relevant to the argument was more important than argumentative moves. In their study, students' transactive reasoning gains were accompanied by reflection as a relevant metalevel (epistemic) process. Planning essays was another relevant metalevel (metacognitive) process accompanying low-structured sensemaking dialogue in a study by Coker Jr and Erwin (2011).

In short, sensemaking tasks, highly represented by LTA Pattern 1 studies, are requisite for advanced argumentation. Making claims clear and precise, justifying claims, and providing reasons for agreeing or disagreeing are common elements across LTA sensemaking tasks.

LTA Pattern 2: High-structured Articulation

Studies belonging to the second pattern (i.e., high-structured articulation) tend to conceive dialogue as a collective negotiation of meanings driven by the critical evaluation of evidence (e.g., Arvaja et al. 2000; Cavagnetto et al. 2010; Ford 2012; Hogan et al. 1999; Kim and Song

2006; Sampson & Clark, 2011). In these studies, argumentation can be used to justify or explain, focusing on the use of evidence-based arguments to draw conclusions (e.g., Baines et al. 2009; Choi et al. 2014; Gillies 2013; Hsu et al. 2015; Kim and Song 2006; Kulatunga et al. 2013; Ryu and Sandoval 2012; Yun and Kim 2015) and/or on the generation, contrast, and evaluation of alternative explanations of the same phenomenon (e.g., Ford 2012; Sampson & Clark, 2011). High-structured dialogue studies focusing on articulation as their major task include operationalizations such as students challenging or questioning each other about evidence (Choi et al. 2014), interpreting and evaluating data (Selcen Guzey and Aranda 2017), or engaging in scientific inquiry (Kim and Song 2006).

When it comes to gains in argument-as-process, Pattern 2 studies tend to place an equal focus on socio-emotional and metalevel processes. For instance, Alexopoulou and Driver (1996) focused on the extent to which students considered and evaluated their own and their peers' assertions instead of simply presenting their views, so that they avoid repetition/circular argument. On top of that reasoning gain, the researchers focused on the socio-emotional process of avoiding tensions and conflict through balancing power in interaction, and on students' willingness and openness to negotiate ideas at a metadialogical level. Similarly, Baines et al. (2009) focused simultaneously on the socio-emotional process of achieving egalitarian student participation through group maintenance versus group blocking, and on the metalevel process of metatalk about the group itself and not about the task. Another example is Ryu and Sandoval (2012) whose main presented student gains were students' listening to each other, at a socio-emotional level, and students' evaluating arguments developing evidentiary norms, at a 'meta' level.

Given the great emphasis on socio-emotional and meta-level processes, as explained above, we might also expect more explicit focus on sophisticated argument products that take alternative points of view into account. However, as shown on Table 3, Pattern 2 studies mainly focus on elaborated (TAP-structure) types of argument. This is probably related to the disciplinary influence that TAP has in science (Erduran et al. 2004), and most of Pattern 2 studies, as we will see later on, come from the science field.

LTA Pattern 3: High-structured Deliberation

When deliberation is the major goal-based task enacted by the studies, a different LTA pattern emerges, always under a highly structured dialogic framing, as with Pattern 2. The difference is that Pattern 3 (high-structured deliberation) studies place the weight not on the articulation between theory and evidence (which is presupposed) but on the articulation between different theories, of which only one can be chosen as the 'best explanation'. Examples of this enactment include considering opposing viewpoints to critique and potentially refine ideas (Golanics and Nussbaum 2008) coming to an agreement on the most sensible solution (Cross et al. 2008) and using argumentation to refocus attention away from personal positions and towards the reasons underlying those positions in light of alternatives (Hsu et al. 2015). The distinction between justificatory and explanatory discourse is also evident in this group of studies (e.g., Asterhan et al. 2012; Noroozi et al. 2013; Oliveira et al. 2015), as with Pattern 2. However, among Pattern 3 studies, attention to metalevel processes is more explicit, including knowledge of the formal qualities of single arguments and of argumentative sequences (Noroozi et al. 2013), regulation of meta-cognition (Hsu et al. 2015), acquisition of argumentation norms as a classroom discourse pattern (Yun and Kim 2015), and willingness to disagree (Kim et al. 2007).

Given the focus on deliberation in Pattern 3 studies, we might again expect more explicit attention to 'dialectic' arguments (integrated arguments taking two contrary perspectives into consideration). However, as shown on Table 3, this is not the case. It seems that dialogic rather than dialectic transactivity is highlighted, which also explains the increased emphasis on meta-level processes. This may be because the majority of these studies take place in small groups, as we will see later on in this section; therefore, a great attention to calibrate a variety of perspectives is necessary.

LTA Pattern 4: Dialectic Deliberation

Finally, studies belonging to the fourth pattern (i.e., dialectic deliberation) explicitly define argumentation as deliberative dialogue (e.g., Felton et al. 2015a, b; Garcia-Mila et al. 2013; Villarroel et al. 2016) with deliberation either being 'purely' dialectical, i.e., focusing on the genuine difference in opinions (Asterhan and Schwarz 2009) or as a path towards consensus building (Berland and Lee 2012). From an instructional framing perspective, the only difference between Pattern 4 and Pattern 3 is that in Pattern 4 studies, disagreement is established as a necessary starting point for the dialogue, whereas in Pattern 3 studies, disagreement may or may not emerge during interaction. This is largely because Pattern 4 studies leverage disagreement to drive gains in counter-argumentation (e.g., Asterhan and Schwarz 2009), rebuttals (e.g., Felton 2004; Kuhn et al. 2016b), and argument-counterargument integration (expressed through revisions, concessions, and compromises) (e.g., Felton et al. 2015b; De Vries et al. 2002; Muller-Mirza et al. 2007; Nussbaum and Edwards 2011). This explicit focus on dialectic deliberation, more often framed in one-to-one rather than small-group settings as illustrated below, leads to a greater focus on the quality of arguments-as-products with the significant majority of Pattern 4 studies focusing on dialectic types of arguments (see Table 3).

LTA Fostering Methods

Considering our whole studies' sample (N = 143), the two main methods for fostering LTA were related to (a) the teacher or coach and (b) the task itself. In total, 72 studies (50%) reported that some type of coach-based method was effective in bringing out any type of gains in students' argumentative dialogue, and 77 studies (54%) did so with task-based methods (32 studies used both coach-based and task-based methods). Among the task-based methods, a majority focused either on the instructional task's structure (e.g., a script for interaction) (37 out of 77 studies), or on an instructional technique adopted as part of the class (e.g., Philosophy for Children) (33 out of 77 studies).

In addition, 41 studies (29%) used some type of material scaffold (visual, online, mixed, or other) to ensure or 'boost' the quality of argumentative dialogue, whereas 28 studies (20%) referred to some type of peers' role, either prescribed or enacted during dialogue, as a relevant fostering factor. Though less common than coach-based and task-based methods, peers' role was an important fostering method either as a variable set a priori or as enacted during interaction. Examples of prescribed peers' role are discussion host (Yun and Kim 2015; Zhang et al. 2016), critical audience (Forman and Ford 2014; Gillies and Haynes 2011), and peer coach (Veerman et al. 2000). Examples of enacted peers' role are leader versus helper (Albe 2008), expert versus follower (Cross et al. 2008), and metacognitive questioner (Gillies 2013).

When considering only the studies belonging to one of the four patterns described above, the distribution of fostering methods per study pattern reveals some further insights regarding how different instructional framings come along with different ways of scaffolding dialogue gains. An interesting observation is that the more we move from low-structured sensemaking to high-structured deliberation, the more concrete the teacher's role becomes, from being a dialogue facilitator to an encourager of transactive reasoning, and from being a thinking guide to a challenger and a critical thinker. Also, the instructor's role in dialectic deliberation vanishes, meaning that the focus is not on what the instructor does or does not, but what the students do as result of the learning environment's guidelines. Alongside this shift, the material supports also move from visual materials (like drafts, graphs, sheets) towards highly structured, often computer-supported, collaborative learning environments. Moreover, the richest study patterns in terms of the variety of scaffolding methods used are Patterns 2 (high-structured articulation) and 3 (high-structured deliberation). Figure 2 presents a comprehensive summary of some representative examples of how LTA dialogue is fostered in each study pattern.

Contextual Aspects of LTA Studies

Overall, the four instructional framing patterns described above seem to define argumentative dialogue in different ways and with different purposes. In this section we will focus on RQ2: In what different ways is argumentative dialogue fostered and/or scaffolded in different contexts? We consider 'context' in the following ways, directly relating to the learner characteristics in our MeDOL framework: (a) the dialogue setting, (b) the subject matter, and (c) students' educational grade.

Study Patterns and Dialogue Setting

To test for an association between the four study patterns and the dialogue setting, we again performed Fisher's exact test, which returned a significant result, $p < .0001$, with a moderate

Study Pattern	Instructor's role/moves	Instructional technique	Task structure/script	Material scaffold
1.Low-structured sensemaking	Open-ended questioning (Molinari & Mameli, 2013; Topping & Trickey, 2007) Probing reasoning (Simon et al., 2008; Oyler, 2019) Facilitator moves/prompts (Kim, 2014; Rezmitskaya et al., 2012; Oyler, 2019)	Collaborative reasoning (Coker & Erwin, 2011; Kim, 2014) Philosophical enquiry/ Philosophy for Children (Rezmitskaya et al., 2012; Topping & Trickey, 2007; Oyler, 2019). Scenarios initiated by puppets used by the teacher (Simon et al., 2008)		Online chat, texts with contradiction or dilemma (Kim, 2014)
2.High-structured articulation	Guiding questions (Cavagnetto et al., 2010; Ruy & Sandoval, 2012) Modelling students' contributions (Ford, 2012; González-Howard et al., 2017; Mercer & Sams, 2006) Probing/challenging students' thinking (Maley et al., 2013; Zhang et al., 2016)	Science writing heuristic (SWH) approach (Cavagnetto et al., 2010; Chen et al., 2016; Choi et al., 2014) SPRinG curriculum (aiming at peer-to-peer interaction) (Baines et al., 2009 Thinking together curriculum (aiming at effective talking and reasoning together) (Mercer & Sams, 2006) IMPROVE technique (focusing on students' metacognitive questioning) (Mevarech & Kramarski, 2003) Inquiry-based instruction (Katchevich et al., 2013; Maley et al., 2013; Moon et al., 2016, 2017; Ruy & Sandoval, 2012)	Critical peer discussion (focusing, exchanging, debating, closing) (Kim & Song, 2005) Guided engagement in scientific writing (Cavagnetto et al., 2009; Chen et al., 2016; Choi et al., 2014) Think-pair-share collaborative script (Hsu et al., 2015) Constructed vs made available diagrams (van Amelsvoort et al., 2007) Task prompt to interpret and evaluate evidence (Arvaja et al., 2000) Task's cognitive demands (Moon et al., 2017)	SWH template (Cavagnetto et al., 2009; Chen et al., 2016; Choi et al., 2014) Online Argumentative Scientific Inquiry System (Hsu et al., 2015) Graphs with information (Ford, 2012; Ruy & Sandoval, 2012) Online chat, texts, diagrams (van Amelsvoort et al., 2007)
3.High-structured deliberation	Discussing evidence and its validity and how it could or could not support claims (Cross, 2009; Evagorou & Osborne, 2013) Encouraging prompts for peer discussions and transactive reasoning (Kim et al., 2007; Noroozi et al., 2013; Yun & Kim, 2015)	Collaborative reasoning (Kim et al., 2007) Writing-based instruction (Corcelles Seuba & Castelló, 2015; Cross, 2009; Forman & Ford, 2014) Structured sociocognitive conflict (Skoumios, 2009)	Questions prompting the use of transactive discourse (Noroozi et al., 2013) Preparatory individual argumentation tasks (Yun & Kim, 2015) Instructions to generate as many reasons as possible crossed with elaborated questions (Golanics & Nussbaum, 2008)	CSCL environments (Evagorou & Osborne, 2013; Hou, 2012; Kim et al., 2007; Noroozi et al., 2013) Planning guide for writing (Corcelles Seuba & Castelló, 2015)
4.Dialectic deliberation		Structured process for arguing, reflecting, arguing again (Kuhn et al., 2008, 2016b)	Arguing to reach consensus versus to persuade (Felton et al., 2015a,b; Garcia-Mila et al., 2013; Gilabert et al., 2013) Argumentative instructions/prompts versus no instructions (Asterhan & Schwarz, 2009; Iordanou et al., 2019; Nussbaum et al., 2005; Theibach et al., 2016; Villarroel et al., 2016)	CSCL environments (De Vries et al., 2002; Harney et al., 2015; Vogel et al., 2016; Weinberger et al., 2013)

Fig. 2 Selective presentation of fostering methods per study pattern

effect size, φc = .545. We can then infer that a first differential contextual aspect of the four study patterns previously identified is the dialogue setting, being: (a) one-to-one, i.e., a peer or dyad of peers arguing 'against' another peer or dyad; (b) small group, i.e., the LTA task relies on student groups to solve/engage with; or (c) whole class, meaning that the teacher guides the dialogue with everyone at a time. Based on analysis of the residuals (Table 4), Pattern 1 studies are mostly associated with whole-class discussions, Pattern 2 and Pattern 3 studies with small-group settings, whereas Pattern 4 studies with one-to-one discussions (as in the case of structured debates).

Study Patterns and Subject Matter

When it comes to the use of a LTA pattern in a specific disciplinary context, we again found a significant association by Fisher's exact test ($p < .0001$), this time between study pattern and subject matter. The effect size was moderate (φc = .415). An analysis of residuals (Table 5) suggests that, the majority of Pattern 2 studies focus on Science (including Maths) and socio-scientific issues (SSI), Pattern 4 mainly focuses on general interest and scientific topics, and Pattern 1 studies focus more on Language/History.

Study Patterns and Educational Grade

Finally, a test of association between the four study patterns and students' grade level produced a significant result using Fisher's exact test (or approached significance, after a Bonferroni corrected alpha level of .01), p = .031, with a moderate effect size, φ_c =.262. As shown on Table 6, Pattern 1 studies were most common at the elementary level, and pattern 4 studies were most common at the university level.

In summary, four study patterns emerged in our sample based on how learning to argue is framed and operationalized as a pedagogical goal:

- Low-structured dialogic sensemaking (Pattern 1) is commonly applied to teacher-mediated whole-class discussions with young children in a variety of disciplinary contexts, with teachers' dialogic moves being used as the main scaffold for students.
- High-structured dialogic articulation (Pattern 2) is associated with TAP structure arguments in small-group settings, particularly in science inquiry where templates and graphs

Table 4 Cross-tabulation of dialogue setting × study pattern (N = 102[a])

		Study pattern				
		Pattern 1	Pattern 2	Pattern 3	Pattern 4	Total
Dialogue setting	One-to-one	1[b]	3[b]	2[b]	18[b]	24
	Small group	4[b]	23[b]	19[b]	9[b]	55
	Whole class	13[b]	3	4	3[b]	23
	Total	18	29	25	30	102

[a] *Of the N = 108 studies belonging to one of the four dialogue study patterns, six applied a mixed dialogue setting*

[b] *Adjusted residuals >2 or <−2*

Table 5 Cross-tabulation of subject matter × study pattern ($N = 108$)

		Study pattern				
		Pattern 1	Pattern 2	Pattern 3	Pattern 4	Total
Subject matter	Science/SSI	6[a]	28[a]	16	11[a]	61
	Language/History	7[a]	1	2	2	12
	General/Other	6	3[a]	7	19[a]	35
	Total	19	32	25	32	108

[a] *Adjusted residuals >2 or <−2*

are used as a primary scaffold for students to construct and compare their scientific explanations.

- High-structured dialogic deliberation (Pattern 3) is associated with a high presence of socio-emotional and meta-level processes, applied mostly with secondary school students particularly with science and SSI topics.
- Dialectic deliberation (Pattern 4) is associated primarily with the production of integrated dialectic arguments in peer-to-peer discussions about general interest issues among older students (adolescents and adults).

Discussion

Despite extensive research, defining the characteristics of effective argumentative dialogue is an open problem, reflected in the complexity and importance of designing argumentative learning environments 'that work' (Berland and McNeill 2010; Bell and Linn 2000; Clark et al. 2007; Jiménez-Aleixandre 2008). Such an enterprise is relatively simple when the goal of argumentative dialogue is some type of measurable conceptual gains in content learning, or performance on a reading/comprehension test. The same is not true when the goal is to promote advances in argumentation itself, either as a *process* or an *outcome*. A common paradox in sociocultural learning emerges: How can learners learn how to argue effectively, when effective engagement in argumentation is a necessary part of such learning? We will address this paradox by referring to two terms commonly used in the literature to refer to effective educational dialogue, namely productivity and constructiveness.

Drawing on Bereiter's ideas, Wells and Arauz (2006) define productive dialogue as one in which 'participants are willing to revise their own opinions as they open-mindedly consider the proposals and arguments of others' and thus 'the common understanding jointly created is

Table 6 Cross-tabulation of educational grade × study pattern ($N = 108$)

		Study pattern				
		Pattern 1	Pattern 2	Pattern 3	Pattern 4	Total
Educational grade	Primary	10[a]	9	4	4	27
	Secondary	8	18	15	17	58
	University	1	5	6	11[a]	23
	Total	19	32	25	32	108

[a] *Adjusted residuals >2 or <−2*

superior to that with which the participants started' (pp. 415–416). Similarly, a dialogue is constructive when it 'literally adds to the (co-)construction or building of something—meaning, understanding, solutions to problems and sometimes knowledge', and when 'it generally contributes in some way to cooperative goal-oriented activity' (Baker 1999; pp. 180–181). In other words, from a productive dialogue perspective, what matters is the production of dialogue moves and sequences that may be considered of a 'high' dialogic quality, such as: open questions, elaboration of previous contributions, reasoned discussion of competing viewpoints, linkage and coordination across contributions, and metacognitive engagement with dialogue (Hennessy et al. 2016; Howe et al. 2019). Furthermore, from the perspective of constructive (argumentative) dialogue, the focus is on the construction of new argumentative and/or content knowledge (e.g., Baker 2009; Noroozi et al. 2013), as part of completing a task with a concrete goal, such as solving a complex problem, addressing an ill-defined issue, or making a decision. When such a goal is reached through a careful calibration of ideas and weighing of points of views, we can even talk about 'productive argumentation' (Andriessen and Schwarz 2009).

This theoretical tripartite of *productive dialogue, constructive argumentative interaction, and productive argumentation* is confirmed by our findings, pointing to the richness of approaches within what can be generally referred to as 'learning to argue' (Muller-Mirza and Perret-Clermont 2009; Von Aufschnaiter et al. 2008). What this review further brings to our knowledge is the instructional effectiveness potential of each one of these theoretical paradigms when applied in concrete contexts. Below we will give a summary of the different instructional framings emerged, discussing the affordances and limitations of each.

Instructional Framing Used to Optimize Classroom Discussions

Under this approach, applied by Pattern 1 studies in our sample, sensemaking is the foundational dialogue goal-task, related to grasping content, framing an authentic question, and co-constructing possible responses. Studies that use instructional framing to optimize classroom discussions tend to focus on dialogue quality per se as an indicator of how thinking and reasoning take place. This characterization of dialogue quality includes exploratory talk (as opposed to presentational or recitation talk) (Brown 2016; Molinari and Mameli 2013), dialogic co-construction of knowledge (Mason 1998), interpretive authority (Chisholm and Loretto 2016), or even presence of arguments (Lee and Majors 2003). Although the focus is not directly on the production of more sophisticated argumentative products, this often is a desired outcome. For example, in Larrain et al. (2014), students' justification of counterarguments and rebuttals is the main indicator of the dialogue's productivity, and it is explicitly fostered by teachers' requests for justifications. Similarly, Sutherland (2006) focuses explicitly on students posing higher-order questions and encouraging critical responses, which are considered a Level 5 argument quality in our coding system (see Table 1). This is made possible, again, through the instructor's explicit orientation towards modelling questions.

When it comes to gains in argument-as-process, Pattern 1 studies in our sample did not typically focus on socio-emotional and meta-level processes as outcomes. And yet, at least conceptually, both processes could play an important role in sensemaking. The collective elaboration and engagement with different views, also known as 'relational agency' (Edwards 2011), is a pre-requisite for social knowledge construction. Learners must not only express their own ideas in dialogue, but also interact with others in ways that invite, interpret, and examine views. Commitment to these processes naturally emerges from a growing epistemic

awareness and meta-level processing of the dialogic process. These are an essential process in any argumentative learning environment, especially when concrete content learning outcomes are expected, because social agency and epistemic agency are so interrelated that one cannot take place without the other. As Miller et al. (2018) remark, for students to be able to act with epistemic agency in the classroom, soliciting and building on each other's knowledge as a resource for learning is a first step in their participatory sensemaking. Additional research exploring these complex relationships in sensemaking dialogue is warranted.

When it comes to fostering methods and conditions related to this type of LTA instructional framing, a preference towards instructors' roles/moves and instructional techniques is observed (Figure 2). It is, therefore, possible that students' LTA gains, among studies following this framing, are largely due to the effectiveness of the teacher in engaging with students' thinking (Leach and Scott 2002; Murphy et al. 2016).

Instructional Framing Used to Optimize Articulation of (Scientific) Contents

Articulation, the main focus of Pattern 2 studies, commonly corresponds, in our review, to collaborative inquiry, i.e., the process in which students engage in collaborative efforts to advance their shared understandings of the phenomenon or issue at hand (Hakkarainen 2003). The passage from sensemaking (content problematization) to articulation, or from Pattern 1 to Pattern 2 study approaches, is marked by a passage from information seeking and elaboration to critique (Chen et al. 2016) or from knowledge-sharing to knowledge-constructing discourse (Fu et al. 2016). Exploratory talk in its full original sense as both construction and critique (Mercer 1995) is what characterizes the dialogical interactions taking place in Pattern 2 articulation studies. But how is this balance between construction and critique achieved? In other words, how can students learn to argue effectively when the focus is on advancing knowledge? One might say that this is possible through transactive reasoning focusing on counterarguments and rebuttals. However, this was the case only in very few studies in our sample (e.g., Ford 2012; Gillies 2016), as the majority of Pattern 2 studies focused on the TAP structure as an indicator of constructive argumentation. This may be because of the focus of scientific inquiry on the construction of scientific explanations, and the TAP structure gives teachers and researchers a (more or less) clear idea of what a scientific explanation looks like (Erduran et al. 2004).

Productive dialogue, for studies focusing on articulation/collaborative inquiry as their main task, is commonly part of interactions where students navigate and negotiate their understandings of phenomena towards more coherent explanations (Baker 1999; Chen et al. 2019). For these moments of collaboration, understood as deep engagement with each other's ideas, to emerge, both argument-as-products and arguments-as-process are necessary: the more constructive students become with their own contributions, the more productive the interaction with each other's contributions becomes (Chi and Menekse 2015). However, what we observe in Pattern 2 studies is a focus on either high-level (4 & 5) arguments-as-products alone, or on a combination of medium-level arguments (2 & 3) with some type of relevant socio-emotional or metalevel processing. We consider this a rich area for inquiry to fill a gap in literature on articulation. As Chen et al. (2019) argue, students' epistemic understanding of argument, especially its plausible and 'uncertain' nature, cannot be separated from their social negotiation processes. What material and visual scaffolds usually do is to reduce such uncertainty, orienting students towards the construction of a theory or solution. Peers' and instructor's role in maintaining uncertainty, through for example the use of critical questions, would be an asset

for this type of studies, as it would increase dialogical transactivity around plausible scientific explanations.

Instructional Framing Used to Optimize the Search for the Best Explanation

Finally, deliberation whether it is framed as small-group interaction (Pattern 3 studies) or one-on-one debates (Pattern 4 studies) focuses on the production of transactive arguments. This includes studies focusing on prompting students to co-construct meanings without necessarily critically contrasting these meanings with each other. In such cases, the learning environment serves as a context of conceptual convergence, understood as a common focus on the construction of shared understanding (Oliveira and Sadler 2008). When this focus on collaboration rather than persuasion also forms an explicit part of the instructional goals for deliberative dialogue, it is often the case that dialectically transactive moves (e.g., claim revision, articulating opposing positions, argument-counterargument integration) emerge in students' discourse (Berland and Lee 2012; Nussbaum 2002; Nussbaum and Edwards 2011; Vogel et al. 2016). This is explained by the principles of coalescent argumentation (Gilbert 2013) according to which collaboration and confrontation are both aspects of a balanced process of critical questioning and coalescing arguments.

Nonetheless, it is also often the case that argumentative dialogue is framed as a process of committing to a position and proving its validity against an opposing position. However, even in this context of adversarial argumentation, a focus on individual reflection on the arguments produced during interaction is often present, either as part of the instructional prompts or as part of the computer-supported collaborative learning environment in which argumentation takes place. This finding is in line with recent recommendations for more research on the role and nature of reflection as a mediating process in students' learning to argue (Iordanou and Constantinou 2014). That said, when students are asked to deliberate in a context of highly divergent positions, the focus must not be on divergence per se, but on how and why these positions differ from each other in terms of their evidence quality and relevance (Macagno 2016, 2019) and how argumentative talk is productively employed from part of the students as required or prompted by the activity design (Schwarz 2009). That latter is not always guaranteed, calling for additional research into the types of instructional prompts and dialogic moves used to support student interaction. This might include, for example, further research into the metacognitive questions asked by teachers, as they move from group to group or from dyad to dyad, which is an aspect studied in depth by the Pattern 1 studies, mostly focusing on whole-group discussions. In addition to the mediatory teacher's role, peers' role can also be further looked at, as recent research suggests that placing the focus on affordances for peer-to-peer counter-argumentation and rebuttal using relevant evidence is an effective strategy for learning to argue (see, for example, Larrain et al. 2019).

Study Limitations

As with meta-analyses (Glass 1976), the integrative review bears the limitation of shifting away from the original units of analysis in the studies reviewed to treating the studies themselves as the unit of analysis. Doing so runs the risk of collapsing disparate findings across settings in the interest of seeing larger patterns in the literature. That being said, we have mitigated this problem by focusing not on the conclusions drawn by each study, but instead on the approach taken by the researchers. Nonetheless, we recognize (per Jackson 1980) that we

have had to make inferences about how researchers frame learning to argue, rather than drawing on explicit statements made by the authors themselves (particularly in the field of educational dialogue, where learning to argue is not necessarily positioned as the goal of dialogue).

A second limitation is that we cannot use our findings to infer which of the studies' characteristics caused the results reported. Whatever statistical results were reported must be seen as a meta-analytical approach of associations of study characteristics coded a posteriori, and not expressing any causal relationship between units of analysis originally used in the studies we reviewed. A final limitation regards the number and type of studies selected to be reviewed. As Jackson (1980) again would point out, the selected studies of a review should be considered only a sample of the phenomenon being studied. Though we have followed recommended practices in sampling from the literature, selecting studies according to their relevance to the phenomenon under investigation (i.e., learning to argue), we recognize that we have not addressed the whole population of studies in the fields of educational dialogue and argumentation, as that would be beyond the scope of this paper and would preclude a focused analysis of the literature.

Conclusion

Constructing learning environments that promote learning to argue (LTA) through dialogue has been characterized as a complex enterprise (Clark et al. 2007; Jiménez-Aleixandre 2008), without sufficiently addressing the reasons for this complexity. We focus on instructional framing as a way to look at the LTA literature, perceiving as such: the methods and approaches implemented by researchers to frame educational dialogue, the dialogue goals pursued as part of the LTA activity, the noted gains in terms of LTA, and the learners' characteristics. Our analysis of the reviewed studies revealed a continuum in the LTA literature, with four different concentrations in empirical studies, namely: engaging in low-structured sensemaking, engaging in high-structured articulation, engaging in high-structured deliberation, and engaging in dialectic deliberation. Each research focus has its own affordances and limitations as discussed above. Interestingly enough, what emerges as a limitation for one trend of research is covered by another, pointing out to the complementarity of existing instructional approaches aiming, either directly or indirectly, at learning to argue.

An implication of the above is that when it comes to teaching students how to argue, all four approaches need to be considered to fully leverage the benefits of engaging in argumentative dialogue. The complementary patterns we have identified clears the way for a better understanding of pedagogical content knowledge in argumentative dialogue, in the interest of designing more effective and purposeful dialogue-based and dialogue-oriented learning environments.

Funding This work is funded by national funds through the FCT – Fundação para a Ciência e a Tecnologia, I. P., under the strategic project UIDB/00183/2020, the Norma Transitória –DL 57/2016/CP1453/CT0066, and the "METACARE: Evidence-based Metaphors for Diabetes Care" Project– PTDC/FER-FIL/28278/2017. The authors would also like to acknowledge the European Commission Horizon 2020 (H2020) project "DIALLS: Dialogue and argumentation for cultural literacy learning in schools" (No 770045), as part of which this study was developed.

Data Availability Data will be available upon request.

Declarations

Conflict of Interest The authors declare no competing interests.

References

Albe, V. (2008). When scientific knowledge, daily life experience, epistemological and social considerations intersect: Students' argumentation in group discussions on a socio-scientific issue. *Research in Science Education, 38*(1), 67–90. https://doi.org/10.1007/s11165-007-9040-2

Alexander, R. J. (2017). *Towards dialogic teaching: Rethinking classroom talk* (5th ed.). Dialogos.

Andriessen, J., & Baker, M. J. (2014). Arguing to learn. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (2nd ed., pp. 439–460). Cambridge University Press.

Andriessen, J. E., & Schwarz, B. B. (2009). Argumentative design. In N. Muller-Mirza & A. N. Perret-Clermont (Eds.), *Argumentation and education: Theoretical foundations and practices* (pp. 145–174). Springer.

Andriessen, J., Baker, M., & van der Puil, C. (2011). Socio-cognitive tension in collaborative working relations. In S. Ludwigsen, A. Lund, I. Rasmussen, & R. Säljö (Eds.), *Learning across sites: New tools, infrastructures and practices* (pp. 222–242). Routledge.

Arvaja, M., Häkkinen, P., Eteläpelto, A., & Rasku-Puttonen, H. (2000). Collaborative processes during report writing of a science learning project: The nature of discourse as a function of task requirements. *European Journal of Psychology of Education, 15*(4), 455–466. https://doi.org/10.1007/bf03172987

Asterhan, C. (2013). Epistemic and interpersonal dimensions of peer argumentation. In M. Baker, J. Andriessen, & S. Järvelä (Eds.), *Affective learning together* (pp. 251–271). Routledge.

Asterhan, C. S., & Schwarz, B. B. (2007). The effects of monological and dialogical argumentation on concept learning in evolutionary theory. *Journal of Educational Psychology, 99*(3), 626–639. https://doi.org/10.1037/0022-0663.99.3.626

Asterhan, C. S., & Schwarz, B. B. (2009). Argumentation and explanation in conceptual change: Indications from protocol analyses of peer-to-peer dialog. *Cognitive Science, 33*(3), 374–400. https://doi.org/10.1111/j.1551-6709.2009.01017.x

Asterhan, C. S., & Schwarz, B. B. (2016). Argumentation for learning: Well-trodden paths and unexplored territories. *Educational Psychologist, 51*(2), 164–187. https://doi.org/10.1080/00461520.2016.1155458

Asterhan, C. S., Schwarz, B. B., & Gil, J. (2012). Small-group, computer-mediated argumentation in middle-school classrooms: The effects of gender and different types of online teacher guidance. *British Journal of Educational Psychology, 82*(3), 375–397. https://doi.org/10.1111/j.2044-8279.2011.02030.x

Asterhan, C. S., Howe, C., Lefstein, A., Matusov, E., & Reznitskaya, A. (2020). Controversies and consensus in research on dialogic teaching and learning. *Dialogic Pedagogy, 8*. https://doi.org/10.5195/dpj.2020.312

Atwood, S., Turnbull, W., & Carpendale, J. I. (2010). The construction of knowledge in classroom talk. *Journal of the Learning Sciences, 19*(3), 358–402. https://doi.org/10.1080/10508406.2010.481013

Aukerman, M., Martin, P. C., Gargani, J., & McCallum, R. D. (2016). A randomized control trial of Shared Evaluation Pedagogy: The near-term and long-term impact of dialogically organized reading instruction. *L1 Educational Studies in Language and Literature, 16*, 1-26. 10.17239/L1ESLL-2016.16.02.02

Baines, E., Rubie-Davies, C., & Blatchford, P. (2009). Improving pupil group work interaction and dialogue in primary classrooms: Results from a year-long intervention study. *Cambridge Journal of Education, 39*(1), 95–117. https://doi.org/10.1080/03057640802701960

Baker, M. J. (1999). Argumentation and constructive interaction. In P. Coirier & J. Andriessen (Eds.), *Foundations of argumentative text processing* (pp. 179–202). University of Amsterdam Press. https://doi.org/10.1007/978-94-017-0781-7_3

Baker, M. (2009). Argumentative interactions and the social construction of knowledge. In N. Muller-Mirza & A. N. Perret-Clermont (Eds.), *Argumentation and education: Theoretical foundations and practices* (pp. 127–144). Springer.

Beasley, T. M., & Schumacker, R. E. (1995). Multiple regression approach to analyzing contingency tables: Post hoc and planned comparison procedures. *Journal of Experimental Education, 64*(1), 79–93. https://doi.org/10.1080/00220973.1995.9943797

Bell, P., & Linn, M. C. (2000). Scientific arguments as learning artifacts: Designing for learning from the web with KIE. *International Journal of Science Education, 22*(8), 797–817. https://doi.org/10.1080/095006900412284

Berkowitz, M. W., & Gibbs, J. C. (1983). Measuring the developmental features of moral discussion. *Merrill-Palmer Quarterly (1982-), 29*, 399–410.

Berland, L. K., & Lee, V. R. (2012). In pursuit of consensus: Disagreement and legitimization during small-group argumentation. *International Journal of Science Education, 34*(12), 1857–1882. https://doi.org/10.1080/09500693.2011.645086

Berland, L. K., & McNeill, K. L. (2010). A learning progression for scientific argumentation: Understanding student work and designing supportive instructional contexts. *Science Education, 94*(5), 765–793. https://doi.org/10.1002/sce.20402

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

Berland, L. K., & Reiser, B. J. (2011). Classroom communities' adaptations of the practice of scientific argumentation. *Science Education, 95*(2), 191–216. https://doi.org/10.1002/sce.20420

Billig, M. (1987). *Arguing and thinking: A rhetorical approach to social psychology*. Cambridge University Press.

Brown, A. C. (2016). Classroom community and discourse: How argumentation emerges during a Socratic circle. *Dialogic Pedagogy: An International Online Journal*, *4*. https://doi.org/10.5195/dpj.2016.160

Cavagnetto, A. R. (2010). Argument to foster scientific literacy: A review of argument interventions in K–12 science contexts. *Review of Educational Research, 80*(3), 336–371. https://doi.org/10.3102/0034654310376953

Cavagnetto, A., Hand, B. M., & Norton-Meier, L. (2010). The nature of elementary student science discourse in the context of the science writing heuristic approach. *International Journal of Science Education, 32*(4), 427–449. https://doi.org/10.1080/09500690802627277

Chen, Y. C., Park, S., & Hand, B. (2016). Examining the use of talk and writing for students' development of scientific conceptual knowledge through constructing and critiquing arguments. *Cognition and Instruction, 34*(2), 100–147. https://doi.org/10.1080/07370008.2016.1145120

Chen, Y. C., Benus, M. J., & Hernandez, J. (2019). Managing uncertainty in scientific argumentation. *Science Education, 103*(5), 1235–1276. https://doi.org/10.1002/sce.21527

Chi, M. T., & Menekse, M. (2015). Dialogue patterns in peer collaboration that promote learning. In L. Resnick, C. Asterhan, & S. Clarke (Eds.), *Socializing intelligence through academic talk and dialogue* (pp. 263–274). American Educational Research Association.

Chin, C., & Teou, L. Y. (2009). Using concept cartoons in formative assessment: Scaffolding students' argumentation. *International Journal of Science Education, 31*(10), 1307–1332. https://doi.org/10.1080/09500690801953179

Chinn, C. A., & Clark, D. B. (2013). Learning through collaborative argumentation. In C. E. Hmelo-Silver, C. A. Chinn, C. K. K. Chan, & A. M. O'Donnell (Eds.), *International handbook of collaborative learning* (pp. 314–332). Taylor & Francis.

Chinn, C. A., Buckland, L. A., & Samarapungavan, A. L. A. (2011). Expanding the dimensions of epistemic cognition: Arguments from philosophy and psychology. *Educational Psychologist, 46*(3), 141–167. https://doi.org/10.1080/00461520.2011.587722

Chisholm, J. S., & Loretto, A. J. (2016). Tensioning interpretive authority during dialogic discussions of literature. *L1 Educational Studies in Language and Literature, 16*, 1–32. https://doi.org/10.17239/L1ESLL-2016.16.02.04

Chiu, M. M. (2008). Effects of argumentation on group micro-creativity: Statistical discourse analyses of algebra students' collaborative problem solving. *Contemporary Educational Psychology, 33*(3), 382–402. https://doi.org/10.1016/j.cedpsych.2008.05.001

Choi, A., Hand, B., & Norton-Meier, L. (2014). Grade 5 students' online argumentation about their in-class inquiry investigations. *Research in Science Education, 44*(2), 267–287. https://doi.org/10.1007/s11165-013-9384-8

Clarà, M. (2021). Conceptually driven inquiry: addressing the tension between dialogicity and teleology in dialogic approaches to classroom talk. *Educational Review.*, 1–20. https://doi.org/10.1080/00131911.2021.1923462

Clark, D. B., Sampson, V., Weinberger, A., & Erkens, G. (2007). Analytic frameworks for assessing dialogic argumentation in online learning environments. *Educational Psychology Review, 19*(3), 343–374. https://doi.org/10.1007/s10648-007-9050-7

Coker Jr., D. L., & Erwin, E. (2011). Teaching academic argument in an urban middle school: A case study of two approaches. *Urban Education, 46*(2), 120–140. https://doi.org/10.1177/0042085910377426

Corcelles Seuba, M., & Castelló, M. (2017). Learning philosophical thinking through collaborative writing in secondary education. *Journal of Writing Research, 7*(1), 157–199. https://doi.org/10.17239/jowr-2015.07.01.07

Cross, D., Taasoobshirazi, G., Hendricks, S., & Hickey, D. T. (2008). Argumentation: A strategy for improving achievement and revealing scientific identities. *International Journal of Science Education, 30*(6), 837–861. https://doi.org/10.1080/09500690701411567

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

Damico, J., & Rosaen, C. L. (2009). Creating epistemological pathways to a critical citizenry: Examination of a fifth-grade discussion of freedom. *Teachers College Record, 111*(5), 1163–1194.

De Vries, E., Lund, K., & Baker, M. (2002). Computer-mediated epistemic dialogue: Explanation and argumentation as vehicles for understanding scientific notions. *Journal of the Learning Sciences, 11*(1), 63–103. https://doi.org/10.1207/s15327809jls1101_3

Dougherty, D., Borrelli, L., Munir, K., & O'Sullivan, A. (2000). Systems of organizational sensemaking for sustained product innovation. *Journal of Engineering and Technology Management, 17*(3-4), 321–355. https://doi.org/10.1016/s0923-4748(00)00028-x

Edwards, A. (2011). Building common knowledge at the boundaries between professional practices: Relational agency and relational expertise in systems of distributed expertise. *International Journal of Educational Research, 50*(1), 33–39. https://doi.org/10.1016/j.ijer.2011.04.007

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*(6), 915–933. https://doi.org/10.1002/sce.20012

Evagorou, M., & Osborne, J. (2013). Exploring young students' collaborative argumentation within a socioscientific issue. *Journal of Research in Science Teaching, 50*(2), 209–237. https://doi.org/10.1002/tea.21076

Felton, M. K. (2004). The development of discourse strategies in adolescent argumentation. *Cognitive Development, 19*(1), 35–52. https://doi.org/10.1016/j.cogdev.2003.09.001

Felton, M., & Kuhn, D. (2001). The development of argumentive discourse skill. *Discourse Processes, 32*(2-3), 135–153. https://doi.org/10.1080/0163853X.2001.9651595

Felton, M., Crowell, A., & Liu, T. (2015a). Arguing to agree: Mitigating my-side bias through consensus-seeking dialogue. *Written Communication, 32*(3), 317–331. https://doi.org/10.1177/0741088315590788

Felton, M., Garcia-Mila, M., Villarroel, C., & Gilabert, S. (2015b). Arguing collaboratively: Argumentative discourse types and their potential for knowledge building. *British Journal of Educational Psychology, 85*(3), 372–386. https://doi.org/10.1111/bjep.12078

Ford, M. (2008). Disciplinary authority and accountability in scientific practice and learning. *Science Education, 92*(3), 404–423. https://doi.org/10.1002/sce.20263

Ford, M. J. (2012). A dialogic account of sense-making in scientific argumentation and reasoning. *Cognition and Instruction, 30*(3), 207–245. https://doi.org/10.1080/07370008.2012.689383

Ford, M. J., & Wargo, B. M. (2012). Dialogic framing of scientific content for conceptual and epistemic understanding. *Science Education, 96*(3), 369–391. https://doi.org/10.1002/sce.20482

Forman, E. A., & Ford, M. J. (2014). Authority and accountability in light of disciplinary practices in science. *International Journal of Educational Research, 64*, 199–210. https://doi.org/10.1016/j.ijer.2013.07.009

Forman, E. A., Ramirez-DelToro, V., Brown, L., & Passmore, C. (2017). Discursive strategies that foster an epistemic community for argument in a biology classroom. *Learning and Instruction, 48*, 32–39. https://doi.org/10.1016/j.learninstruc.2016.08.005

Frijters, S., ten Dam, G., & Rijlaarsdam, G. (2008). Effects of dialogic learning on value-loaded critical thinking. *Learning and Instruction, 18*(1), 66–82. https://doi.org/10.1016/j.learninstruc.2006.11.001

Fu, E. L., van Aalst, J., & Chan, C. K. (2016). Toward a classification of discourse patterns in asynchronous online discussions. *International Journal of Computer-Supported Collaborative Learning, 11*(4), 441–478. https://doi.org/10.1007/s11412-016-9245-3

Garcia-Mila, M., Gilabert, S., Erduran, S., & Felton, M. (2013). The effect of argumentative task goal on the quality of argumentative discourse. *Science Education, 97*(4), 497–523. https://doi.org/10.1002/sce.21057

Gilabert, S., Garcia-Mila, M., & Felton, M. K. (2013). The effect of task instructions on students' use of repetition in argumentative discourse. *International Journal of Science Education, 35*(17), 2857–2878. https://doi.org/10.1080/09500693.2012.663191

Gilbert, M. A. (2004). Emotion, argumentation and informal logic. *Informal Logic, 24*(3), 245–264. https://doi.org/10.22329/il.v24i3.2147

Gilbert, M. A. (2013). *Coalescent argumentation*. Routledge.

Gillies, R. M. (2013). Productive academic talk during inquiry-based science. *Pedagogies: An International Journal, 8*(2), 126–142. https://doi.org/10.1080/1554480x.2013.767770

Gillies, R. M. (2016). Dialogic interactions in the cooperative classroom. *International Journal of Educational Research, 76*, 178–189. https://doi.org/10.1016/j.ijer.2015.02.009

Gillies, R. M., & Haynes, M. (2011). Increasing explanatory behaviour, problem-solving, and reasoning within classes using cooperative group work. *Instructional Science, 39*(3), 349–366. https://doi.org/10.1007/s11251-010-9130-9

Glass, G. V. (1976). Primary, secondary, and meta-analysis of research. *Educational Researcher, 5*, 3–8.

Golanics, J. D., & Nussbaum, E. M. (2008). Enhancing online collaborative argumentation through question elaboration and goal instructions. *Journal of Computer Assisted Learning, 24*(3), 167–180. https://doi.org/10.1111/j.1365-2729.2007.00251.x

González-Howard, M., McNeill, K. L., Marco-Bujosa, L. M., & Proctor, C. P. (2017). 'Does it answer the question or is it French fries?': An exploration of language supports for scientific argumentation. *International Journal of Science Education, 39*(5), 528–547. https://doi.org/10.1080/09500693.2017.1294785

Hakkarainen, K. (2003). Progressive inquiry in a computer-supported biology class. *Journal of Research in Science Teaching, 40*(10), 1072–1088. https://doi.org/10.1002/tea.10121

Hand, B., Chen, Y. C., & Suh, J. K. (2021). Does a knowledge generation approach to learning benefit students? A systematic review of research on the science writing heuristic approach. *Educational Psychology Review, 33*(2), 535–577. https://doi.org/10.1007/s10648-020-09550-0

Harney, O. M., Hogan, M. J., Broome, B., Hall, T., & Ryan, C. (2015). Investigating the effects of prompts on argumentation style, consensus and perceived efficacy in collaborative learning. *International Journal of Computer-Supported Collaborative Learning, 10*(4), 367–394. https://doi.org/10.1007/s11412-015-9223-1

Hennessy, S., Rojas-Drummond, S., Higham, R., Márquez, A. M., Maine, F., Ríos, R. M., García-Carrión, R., Torreblanca, O., & Barrera, M. J. (2016). Developing a coding scheme for analysing classroom dialogue across educational contexts. *Learning, Culture and Social Interaction, 9*, 16–44. https://doi.org/10.1016/j.lcsi.2015.12.001

Hoadley, U. (2006). Analysing pedagogy: The problem of framing. *Journal of Education, 40*(1), 15–34.

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/s1532690xci1704_2

Howe, C., Hennessy, S., Mercer, N., Vrikki, M., & Wheatley, L. (2019). Teacher-student dialogue during classroom teaching: Does it really impact on student outcomes? *Journal of the Learning Sciences, 28*(4-5), 462–512. https://doi.org/10.1080/10508406.2019.1573730

Hsu, P. S., Van Dyke, M., Chen, Y., & Smith, T. J. (2015). The effect of a graph-oriented computer-assisted project-based learning environment on argumentation skills. *Journal of Computer Assisted Learning, 31*(1), 32–58. https://doi.org/10.1111/jcal.12080

Iordanou, K., & Constantinou, C. P. (2014). Developing pre-service teachers' evidence-based argumentation skills on socio-scientific issues. *Learning and Instruction, 34*, 42–57. https://doi.org/10.1016/j.learninstruc.2014.07.004

Iordanou, K., Kendeou, P., & Beker, K. (2016). Argumentative reasoning. In J. A. Greene, W. A. Sandoval, & I. Braten (Eds.), *Handbook of epistemic cognition* (pp. 39–53). Routledge.

Iordanou, K., Kuhn, D., Matos, F., Shi, Y., & Hemberger, L. (2019). Learning by arguing. *Learning and Instruction*,. Online first, *63*, 101207. https://doi.org/10.1016/j.learninstruc.2019.05.004

Jackson, G. B. (1980). Methods for integrative reviews. *Review of Educational Research, 50*(3), 438–460.

Jadallah, M., Anderson, R. C., Nguyen-Jahiel, K., Miller, B. W., Kim, I. H., Kuo, L. J., Dong, T., & Wu, X. (2011). Influence of a teacher's scaffolding moves during child-led small-group discussions. *American Educational Research Journal, 48*(1), 194–230. https://doi.org/10.3102/0002831210371498

Jermann, P., & Dillenbourg, P. (2003). Elaborating new arguments through a CSCL script. In J. Andriessen, M. Baker, & D. Suthers (Eds.), *Arguing to learn: Confronting cognitions in computer-supported collaborative learning environments* (pp. 205–226). Springer.

Jiménez-Aleixandre, M. P. (2002). Knowledge producers or knowledge consumers? Argumentation and decision making about environmental management. *International Journal of Science Education, 24*(11), 1171–1190. https://doi.org/10.1080/09500690210134857

Jiménez-Aleixandre, M.-P. (2008). Designing argumentation learning environments. In S. Erduran & M.-P. Jiménez-Aleixandre (Eds.), *Argumentation in science education* (pp. 91–116). Springer.
Katchevich, D., Hofstein, A., & Mamlok-Naaman, R. (2013). Argumentation in the chemistry laboratory: Inquiry and confirmatory experiments. *Research in Science Education, 43*(1), 317–345. https://doi.org/10.1007/s11165-011-9267-9
Keefer, M. W., Zeitz, C. M., & Resnick, L. B. (2000). Judging the quality of peer-led student dialogues. *Cognition and Instruction, 18*(1), 53–81. https://doi.org/10.1207/s1532690xci1801_03
Kim, H., & Song, J. (2006). The features of peer argumentation in middle school students' scientific inquiry. *Research in Science Education, 36*(3), 211–233. https://doi.org/10.1007/s11165-005-9005-2
Kim, M. Y., & Wilkinson, I. A. (2019). What is dialogic teaching? Constructing, deconstructing, and reconstructing a pedagogy of classroom talk. *Learning, Culture and Social Interaction, 21*, 70–86. https://doi.org/10.1016/j.lcsi.2019.02.003
Kim, I. H. (2014). Development of reasoning skills through participation in collaborative synchronous online discussions. *Interactive Learning Environments, 22*(4), 467–484. https://doi.org/10.1080/10494820.2012.680970
Kim, I.-H., Anderson, R. C., Nguyen-Jahiel, K., & Archodidou, A. (2007). Discourse patterns during children's collaborative online discussions. *Journal of the Learning Sciences, 16*(3), 333–370. https://doi.org/10.1080/10508400701413419
Klein, P. D., Haug, K. N., & Bildfell, A. (2019). Writing to learn. In S. Graham, C. A. McArthur, & M. Hebert (Eds.), *Best practices in writing instruction* (3rd ed., pp. 162–184). The Guilford Press.
Koschmann, T. (1999). Towards a dialogic theory of learning: Bakhtin's contribution to understanding learning in settings of collaboration. In C. Hoadley & J. Roschelle (Eds.), *Proceedings of the Computer Support for Collaborative Learning (CSCL) 1999 Conference* (pp. 308–313). Lawrence Erlbaum Associates.
Krabbe, E. C. W. (2003). Metadialogues. In F. H. van Eemeren, J. A. Blair, C. A. Willard, & A. F. Snoeck Henkemans (Eds.), *Anyone who has a view: Theoretical contributions to the study of argumentation* (pp. 83–90). Kluwer.
Kruger, A. C. (1993). Peer collaboration: Conflict, cooperation, or both? *Social Development, 2*(3), 165–182. https://doi.org/10.1111/j.1467-9507.1993.tb00012.x
Kuhn, D. (1999). A developmental model of critical thinking. *Educational Researcher, 28*(2), 16–46. https://doi.org/10.3102/0013189x028002016
Kuhn, D. (2015). Thinking together and alone. *Educational Researcher, 44*(1), 46–53. https://doi.org/10.3102/0013189x15569530
Kuhn, D. (2018). *Building our best future: Thinking critically about ourselves and our world*. 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*, 545–552. https://doi.org/10.1177/0956797611402512
Kuhn, D., Goh, W., Iordanou, K., & Shaenfield, D. (2008). Arguing on the computer: A microgenetic study of developing argument skills in a computer-supported environment. *Child Development, 79*(5), 1310–1328. https://doi.org/10.1111/j.1467-8624.2008.01190.x
Kuhn, D., & Udell, W. (2003). The development of argument skills. *Child Development, 74*(5), 1245–1260. https://doi.org/10.1111/1467-8624.00605
Kuhn, D., & Udell, W. (2007). Coordinating own and other perspectives in argument. *Thinking and Reasoning, 13*(2), 90–104. https://doi.org/10.1080/13546780600625447
Kuhn, D., & Zillmer, N. (2015). Developing norms of discourse. In L. Resnick, C. Asterhan, & S. Clarke (Eds.), *Socializing intelligence through academic talk and dialogue* (pp. 77–86). American Educational Research Association. https://doi.org/10.3102/978-0-935302-43-1_6
Kuhn, D., Shaw, V., & Felton, M. (1997). Effects of dyadic interaction on argumentive reasoning. *Cognition and Instruction, 15*(3), 287–315. https://doi.org/10.1207/s1532690xci1503_1
Kuhn, D., Zillmer, N., Crowell, A., & Zavala, J. (2013). Developing norms of argumentation: Metacognitive, epistemological, and social dimensions of developing argumentive competence. *Cognition and Instruction, 31*(4), 456–496. https://doi.org/10.1080/07370008.2013.830618
Kuhn, D., Hemberger, L., & Khait, V. (2016a). *Argue with me: Developing thinking and writing through dialog*. Routledge.
Kuhn, D., Hemberger, L., & Khait, V. (2016b). Tracing the development of argumentive writing in a discourse-rich context. *Written Communication, 33*(1), 92–121. https://doi.org/10.1177/0741088315617157
Kulatunga, U., Moog, R. S., & Lewis, J. E. (2013). Argumentation and participation patterns in general chemistry peer-led sessions. *Journal of Research in Science Teaching, 50*(10), 1207–1231.

Langer-Osuna, J. M. (2015). From getting 'fired' to becoming a collaborator: A case of the coconstruction of identity and engagement in a project-based mathematics classroom. *Journal of the Learning Sciences, 24*(1), 53–92. https://doi.org/10.1080/10508406.2014.944643

Langer-Osuna, J. M., & Avalos, M. A. (2015). 'I'm trying to figure this out. Why don't you come up here?': Heterogeneous talk and dialogic space in a mathematics discussion. *ZDM, 47*(7), 1313–1322. https://doi.org/10.1007/s11858-015-0735-y

Larrain, A., Howe, C., & Cerda, J. (2014). Argumentation in whole-class teaching and science learning. *Psykhe, 23*(2), 1–15. https://doi.org/10.7764/psykhe.23.2.712

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, 37*(4), 453–482. https://doi.org/10.1080/07370008.2019.1627360

Leach, J., & Scott, P. (2002). Designing and evaluating science teaching sequences: An approach drawing upon the concept of learning demand and a social constructivist perspective on learning. *Studies in Science Education, 38*, 115–142.

Lee, C. D., & Majors, Y. J. (2003). 'Heading up the street': Localised opportunities for shared constructions of knowledge. *Pedagogy, Culture and Society, 11*(1), 49–68. https://doi.org/10.1080/14681360300200160

Lee, S., Kang, E., & Kim, H. B. (2015). Exploring the impact of students' learning approach on collaborative group modeling of blood circulation. *Journal of Science Education and Technology, 24*(2–3), 234–255. https://doi.org/10.1007/s10956-014-9509-5.

Leitão, S. (2000). The potential of argument in knowledge building. *Human Development, 43*(6), 332–360. https://doi.org/10.1159/000022695

Lipman, M. (2003). *Thinking in education*. Cambridge University Press.

Littleton, K., & Mercer, N. (2013). *Interthinking: Putting talk to work*. Routledge.

Macagno, F. (2016). Argument relevance and structure. Assessing and developing students' uses of evidence. *International Journal of Educational Research, 79*, 180–194. https://doi.org/10.1016/j.ijer.2016.07.002

Macagno, F. (2019). Coding relevance. *Learning, culture, and social interaction*. Online first, 100349. https://doi.org/10.1016/j.lcsi.2019.100349

Macagno, F., Mayweg-Paus, E., & Kuhn, D. (2015). Argumentation theory in education studies: Coding and improving students' argumentative strategies. *Topoi, 34*(2), 523–537. https://doi.org/10.1007/s11245-014-9271-6

Maley, T., Stoll, W., & Demir, K. (2013). Seeing an old lab in a new light: Transforming a traditional optics lab into full guided inquiry. *The Physics Teacher, 51*(6), 368–371. https://doi.org/10.1119/1.4818379

Mason, L. (1998). Sharing cognition to construct scientific knowledge in school context: The role of oral and written discourse. *Instructional Science, 26*(5), 359–389. https://doi.org/10.1023/a:1003103213786

McNeill, K. L., & Pimentel, D. S. (2010). Scientific discourse in three urban classrooms: The role of the teacher in engaging high school students in argumentation. *Science Education, 94*(2), 203–229. https://doi.org/10.1002/sce.20364

Mercer, N. (1995). *The guided construction of knowledge: Talk amongst teachers and learners*. Multilingual matters.

Mercer, N. (2004). Sociocultural discourse analysis. *Journal of Applied Linguistics, 1*(2), 137–168. https://doi.org/10.1558/japl.v1i2.137

Mercer, N., & Howe, C. (2012). Explaining the dialogic processes of teaching and learning: The value and potential of sociocultural theory. *Learning, Culture and Social Interaction, 1*(1), 12–21. https://doi.org/10.1016/j.lcsi.2012.03.001

Mercer, N., & Sams, C. (2006). Teaching children how to use language to solve maths problems. *Language and Education, 20*(6), 507–528. https://doi.org/10.2167/le678.0

Mevarech, Z. R., & Kramarski, B. (2003). The effects of metacognitive training versus worked-out examples on students' mathematical reasoning. *British Journal of Educational Psychology, 73*(4), 449–471. https://doi.org/10.1348/000709903322591181

Meyer, O., Coyle, D., Halbach, A., Schuck, K., & Ting, T. (2015). A pluriliteracies approach to content and language integrated learning–mapping learner progressions in knowledge construction and meaning-making. *Language, Culture and Curriculum, 28*(1), 41–57. https://doi.org/10.1080/07908318.2014.1000924

Michaels, S., O'Connor, M. C., Hall, M. W., & Resnick, L. (2002). *Accountable talk: Classroom conversation that works*. University of Pittsburgh Press.

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

Micheli, R. (2012). Arguing without trying to persuade? Elements for a non-persuasive definition of argumentation. *Argumentation, 26*(1), 115–126. https://doi.org/10.1007/s10503-011-9240-9

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

Molinari, L., & Mameli, C. (2013). Process quality of classroom discourse: Pupil participation and learning opportunities. *International Journal of Educational Research, 62*, 249–258. https://doi.org/10.1016/j.ijer.2013.05.003

Moon, A., Stanford, C., Cole, R., & Towns, M. (2016). The nature of students' chemical reasoning employed in scientific argumentation in physical chemistry. *Chemistry Education Research and Practice, 17*(2), 353–364. https://doi.org/10.1039/c5rp00207a

Moon, A., Stanford, C., Cole, R., & Towns, M. (2017). Analysis of inquiry materials to explain complexity of chemical reasoning in physical chemistry students' argumentation. *Journal of Research in Science Teaching, 54*(10), 1322–1346. https://doi.org/10.1002/tea.21407

Morrison, G. R., Ross, S. M., Morrison, J. R., & Kalman, H. K. (2019). *Designing effective instruction* (8th ed.). John Wiley and Sons.

Muller-Mirza, N., & Perret-Clermont, A. N. (2009). *Argumentation and education: Theoretical foundations and practices*. Springer.

Muller-Mirza, N., Tartas, V., Perret-Clermont, A. N., & de Pietro, J. F. (2007). Using graphical tools in a phased activity for enhancing dialogical skills: An example with Digalo. *International Journal of Computer-Supported Collaborative Learning, 2*(2), 247–272. https://doi.org/10.1007/s11412-007-9021-5

Murphy, P. K., Firetto, C. M., Wei, L., Li, M., & Croninger, R. M. (2016). What REALLY works: Optimizing classroom discussions to promote comprehension and critical-analytic thinking. *Policy Insights from the Behavioral and Brain Sciences, 3*(1), 27–35. https://doi.org/10.1177/2372732215624215

Nielsen, J. A. (2013). Dialectical features of students' argumentation: A critical review of argumentation studies in science education. *Research in Science Education, 43*(1), 371–393. https://doi.org/10.1007/s11165-011-9266-x

Noroozi, O., Weinberger, A., Biemans, H. J., Mulder, M., & Chizari, M. (2013). Facilitating argumentative knowledge construction through a transactive discussion script in CSCL. *Computers & Education, 61*, 59–76. https://doi.org/10.1016/j.compedu.2012.08.013

Noroozi, O., Kirschner, P. A., Biemans, H. J., & Mulder, M. (2018). Promoting argumentation competence: Extending from first-to second-order scaffolding through adaptive fading. *Educational Psychology Review, 30*(1), 153–176. https://doi.org/10.1007/s10648-017-9400-z

Nussbaum, E. M. (2002). Scaffolding argumentation in the social studies classroom. *The Social Studies, 93*(2), 79–83. https://doi.org/10.1080/00377990209599887

Nussbaum, E. M. (2008). Using argumentation vee diagrams (AVDs) for promoting argument-counterargument integration in reflective writing. *Journal of Educational Psychology, 100*(3), 549–565. https://doi.org/10.1037/0022-0663.100.3.549

Nussbaum, E. M., & Edwards, O. V. (2011). Critical questions and argument stratagems: A framework for enhancing and analyzing students' reasoning practices. *Journal of the Learning Sciences, 20*(3), 443–488. https://doi.org/10.1080/10508406.2011.564567

O'Keefe, D. J. (1992). Two concepts of argument. In W. L. Benoit, D. Hample, & P. Benoit (Eds.), *Readings in argumentation* (pp. 79–90). Foris Publications.

Oliveira, A. W., & Sadler, T. D. (2008). Interactive patterns and conceptual convergence during student collaborations in science. *Journal of Research in Science Teaching, 45*(5), 634–658. https://doi.org/10.1002/tea.20211

Oliveira, D. K. B., Justi, R., & Mendonça, P. C. C. (2015). The use of representations and argumentative and explanatory situations. *International Journal of Science Education, 37*(9), 1402–1435. https://doi.org/10.1080/09500693.2015.1039095

Osborne, J., Simon, S., Christodoulou, A., Howell-Richardson, C., & Richardson, K. (2013). Learning to argue: A study of four schools and their attempt to develop the use of argumentation as a common instructional practice and its impact on students. *Journal of Research in Science Teaching, 50*(3), 315–347. https://doi.org/10.1002/tea.21073

Osborne, J. F., Henderson, J. B., MacPherson, A., Szu, E., Wild, A., & Yao, S. Y. (2016). The development and validation of a learning progression for argumentation in science. *Journal of Research in Science Teaching, 53*(6), 821–846.

Oyler, J. (2019). Exploring teacher contributions to student argumentation quality. *Studia Paedagogica, 24*(4), 173–198. https://doi.org/10.5817/sp2019-4-8

Plantin, C. (2004). On the inseparability of emotion and reason in argumentation. *Amsterdam Studies in the Theory and History of Linguistic Sciences Series, 4*(248), 265–276. https://doi.org/10.1075/cilt.248.18pla

Polo, C., Lund, K., Plantin, C., & Niccolai, G. P. (2016). Group emotions: The social and cognitive functions of emotions in argumentation. *International Journal of Computer-Supported Collaborative Learning, 11*(2), 123–156. https://doi.org/10.1007/s11412-016-9232-8

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.* Wilmington: Vernon Press.

Rapanta, C. (2021). Can teachers implement a student-centered dialogical argumentation method across the curriculum?. *Teaching and Teacher Education, 105*. https://doi.org/10.1016/j.tate.2021.103404

Rapanta, C., Garcia-Mila, M., & Gilabert, S. (2013). What is meant by argumentative competence? An integrative review of methods of analysis and assessment in education. *Review of Educational Research, 83*(4), 483-520. https://doi.org/10.3102/0034654313487606

Rapanta, C., & Macagno, F. (2019). Pragmatics, education and argumentation: Introduction to the special issue. *Learning, Culture and Social Interaction.* https://doi.org/10.1016/j.lcsi.2019.100371

Ravenscroft, A. (2000). Designing argumentation for conceptual development. *Computers & Education, 34*(3-4), 241–255. https://doi.org/10.1016/s0360-1315(99)00048-2

Resnick, L. B., Asterhan, C. S., Clarke, S. N., & Schantz, F. (2018). Next generation research in dialogic learning. In G. E. Hall, L. E. Quinn, & D. M. Gollnick (Eds.), *Wiley handbook of teaching and learning* (pp. 323–338). Wiley.

Reznitskaya, A., Kuo, L. J., Clark, A. M., Miller, B., Jadallah, M., Anderson, R. C., & Nguyen-Jahiel, K. (2009). Collaborative reasoning: A dialogic approach to group discussions. *Cambridge Journal of Education, 39*(1), 29–48. https://doi.org/10.1080/03057640802701952

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

Ryu, S., & Sandoval, W. A. (2012). Improvements to elementary children's epistemic understanding from sustained argumentation. *Science Education, 96*(3), 488–526. https://doi.org/10.1002/sce.21006

Sampson, V., & Clark, D. B. (2011). A comparison of the collaborative scientific argumentation practices of two high and two low performing groups. *Research in Science Education, 41*(1), 63–97. https://doi.org/10.1007/s11165-009-9146-9.

Sandelowski, M., & Barroso, J. (2007). *Handbook for synthesizing qualitative research*. Springer.

Schwarz, B. B. (2009). Argumentation and learning. In N. Muller-Mirza & A. N. Perret-Clermont (Eds.), *Argumentation and education: Theoretical foundations and practices* (pp. 91–126). Springer.

Scott, P. H., Mortimer, E. F., & Aguiar, O. G. (2006). The tension between authoritative and dialogic discourse: A fundamental characteristic of meaning making interactions in high school science lessons. *Science Education, 90*(4), 605–631. https://doi.org/10.1002/sce.20131

Selcen Guzey, S., & Aranda, M. (2017). Student participation in engineering practices and discourse: An exploratory case study. *Journal of Engineering Education, 106*(4), 585–606. https://doi.org/10.1002/jee.20176

Sfard, A. (1998). On two metaphors for learning and the dangers of choosing just one. *Educational Researcher, 27*(2), 4–13. https://doi.org/10.3102/0013189X027002004

Simon, S., Naylor, S., Keogh, B., Maloney, J., & Downing, B. (2008). Puppets promoting engagement and talk in science. *International Journal of Science Education, 30*(9), 1229–1248. https://doi.org/10.1080/09500690701474037

Skoumios, M. (2009). The effect of sociocognitive conflict on students' dialogic argumentation about floating and sinking. *International Journal of Environmental and Science Education, 4*(4), 381–399.

Sutherland, J. (2006). Promoting group talk and higher-order thinking in pupils by 'coaching' secondary English trainee teachers. *Literacy, 40*(2), 106–114. https://doi.org/10.1111/j.1467-9345.2006.00436.x

Thiebach, M., Mayweg-Paus, E., & Jucks, R. (2016). Better to agree or disagree? The role of critical questioning and elaboration in argumentative discourse. *Zeitschrift für Pädagogische Psychologie, 30*(2-3), 133–149. https://doi.org/10.1024/1010-0652/a000174

Topping, K. J., & Trickey, S. (2007). Impact of philosophical enquiry on school students' interactive behaviour. *Thinking Skills and Creativity, 2*(2), 73–84. https://doi.org/10.1016/j.tsc.2007.03.001

Toulmin, S. (1958). *The uses of argument*. Cambridge University Press.

Van Amelsvoort, M., Andriessen, J., & Kanselaar, G. (2007). Representational tools in computer-supported collaborative argumentation-based learning: How dyads work with constructed and inspected argumentative diagrams. *The Journal of the Learning Sciences, 16*(4), 485–521. https://doi.org/10.1080/10508400701524785

van Eemeren, F. H., & Grootendorst, R. (1992). *Argumentation, communication, and fallacies: A pragma-dialectical perspective*. Lawrence Erlbaum Associates, Inc..

Van Eemeren, F. H., & Grootendorst, R. (2003). A pragma-dialectical procedure for a critical discussion. *Argumentation, 17*(4), 365–386. https://doi.org/10.1023/a:1026334218681

Veerman, A. L., Andriessen, J. E., & Kanselaar, G. (2000). Learning through synchronous electronic discussion. *Computers & Education, 34*(3-4), 269–290. https://doi.org/10.1016/S0360-1315(99)00050-0

Villarroel, C., Felton, M., & Garcia-Mila, M. (2016). Arguing against confirmation bias: The effect of argumentative discourse goals on the use of disconfirming evidence in written argument. *International Journal of Educational Research, 79*, 167–179. https://doi.org/10.1016/j.ijer.2016.06.009

Vogel, F., Kollar, I., Ufer, S., Reichersdorfer, E., Reiss, K., & Fischer, F. (2016). Developing argumentation skills in mathematics through computer-supported collaborative learning: The role of transactivity. *Instructional Science, 44*(5), 477–500. https://doi.org/10.1007/s11251-016-9380-2

Von Aufschnaiter, C., Erduran, S., Osborne, J., & Simon, S. (2008). Arguing to learn and learning to argue: Case studies on 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

Walton, D. N. (1989). *Informal logic: A handbook for critical argumentation*. Cambridge University Press.

Walton, D. N. (1998). *The new dialectic: Conversational contexts of argument*. University of Toronto Press.

Walton, D. (2013). *Methods of argumentation*. Cambridge University Press.

Walton, D., & Krabbe, E. C. (1995). *Commitment in dialogue: Basic concepts of interpersonal reasoning*. SUNY press.

Weinberger, A., Marttunen, M., Laurinen, L., & Stegmann, K. (2013). Inducing socio-cognitive conflict in Finnish and German groups of online learners by CSCL script. *International Journal of Computer-Supported Collaborative Learning, 8*(3), 333–349. https://doi.org/10.1007/s11412-013-9173-4

Wells, G. (2007). Semiotic mediation, dialogue and the construction of knowledge. *Human Development, 50*(5), 244–274. https://doi.org/10.1159/000106414

Wells, G., & Arauz, R. M. (2006). Dialogue in the classroom. *The Journal of the Learning Sciences, 15*(3), 379–428. https://doi.org/10.1207/s15327809jls1503_3

Whittemore, R., & Knafl, K. (2005). The integrative review: Updated methodology. *Journal of Advanced Nursing, 52*(5), 546–553. https://doi.org/10.1111/j.1365-2648.2005.03621.x

Wu, H. K., & Krajcik, J. S. (2006). Inscriptional practices in two inquiry-based classrooms: A case study of seventh graders' use of data tables and graphs. *Journal of Research in Science Teaching, 43*(1), 63–95. https://doi.org/10.1002/tea.20092

Yun, S. M., & Kim, H. B. (2015). Changes in students' participation and small group norms in scientific argumentation. *Research in Science Education, 45*(3), 465–484. https://doi.org/10.1007/s11165-014-9432-z

Nussbaum, E. M., Kardash, C. M., & Graham, S. E. (2005). The effects of goal instructions and text on the generation of counterarguments during writing. *Journal of Educational Psychology, 97*(2), 157–169. https://doi.org/10.1037/0022-0663.97.2.157

Zhang, J., Niu, C., Munawar, S., & Anderson, R. C. (2016). What makes a more proficient discussion group in English language learners' classrooms? Influence of teacher talk and student backgrounds. *Research in the Teaching of English*, 183–208.

Zohar, A., & Nemet, F. (2002). Fostering students' knowledge and argumentation skill through dilemmas in human genetics. *Journal of Research in Science Teaching, 39*(1), 35–62. https://doi.org/10.1002/tea.10008