Nera et al. *Research and Practice in Technology Enhanced Learning* (2025) 20:17
https://doi.org/10.58459/rptel.2025.20017

Research and Practice in Technology Enhanced Learning

RESEARCH Free and Open Access

Improving argumentation quality on MOOC discussion forums: does learning to identify components of arguments help?

Kenzo Nera [1] *, Mariane Frenay [2] and Magali Paquot [3]

*Correspondence:
kenzo.nera@ulb.be
Psychological Sciences Research Institute (IPSY),
Institute for Language and Communication (ILC),
Université catholique de Louvain,
Place du Cardinal Mercier, 10 bte L3.05.01, Louvain-la-Neuve,
B-1348, Belgium
Full list of author information is available at the end of the article.

Abstract

Constructive argumentation among learners is integral to effective learning. In the context of Massively Open Online Classes (MOOCs), such peer interactions can only occur in discussion forums, where they often prove to be sparse and of poor quality. To address these challenges, we developed and experimentally tested an intervention ($n_{participants}$ = 110, $n_{arguments}$ = 270) aimed at improving the quality of learners' written arguments in MOOC forums, taking into account MOOC platform constraints (e.g., self-paced participation and the impossibility of providing personalised feedback). In the first chapter of a management MOOC, participants randomly assigned to the experimental group (vs control) were introduced to the formal components of arguments: claims, justifications and qualifications. They were then asked to identify these components in a series of examples. We found no significant impact of this intervention on learners' responses to individual open-ended questions directly following the intervention, or in their contributions to discussion forums. Instead, we observed variation in argument quality based on the specific questions prompting argumentation. Our findings prompt further discussion and exploration of strategies to enhance argumentation quality in MOOC discussion forums.

Keywords: MOOC, Forums, Argumentation, Learning, Toulmin's Model of Arguments

Introduction

Massive Open Online Courses (MOOCs) have been on the rise for the past twenty years and have deeply impacted the practices of universities, professors and learners (Daniel, 2012; Voudoukis & Pagiatakis, 2022). MOOCs enable students from all socioeconomic and geographical backgrounds to access higher education (Weinhardt & Sitzmann, 2019).

However, MOOCs also have important limitations (Waldrop, 2013). In particular, their self-paced and online nature constrains learner-learner and learner-teacher interactions (e.g., López Meneses et al., 2020). Whereas in traditional "in person" learning there are many opportunities for interactions to occur (e.g., asking questions), the only spaces for such interactions in MOOCs are discussion forums. Unfortunately, interactions on MOOC forums have typically been reported to be sparse and of low quality (Galikyan et al., 2021; Li et al., 2020; Tawfik et al., 2017).

Arguing with fellow learners about class content fosters learning (Kuhn, 1993; Schwarz, 2009; see, however, Asterhan & Schwarz, 2016; Wecker & Fischer, 2014). There is therefore every justification for developing ways to improve the quality of interactions on MOOC discussion forums. In this research, we developed and experimentally tested an intervention aimed at improving the quality of learners' arguments in MOOC discussion forums exploring essentially contested concepts, that is, concepts that are controversial in nature (Collier et al., 2006; Gallie, 1956). The goal of the intervention was to introduce learners to the formal components of arguments in accordance with the simplified version of Toulmin's model of arguments (Toulmin, 1958; Weinberger & Fischer, 2006). Finally, we wished to develop an intervention adapted to the constraints of platforms hosting MOOCs, and specifically edX (e.g., self-paced participation).

The benefits of argumentation for learning

Argumentation can be defined as a "verbal and social activity of reason aimed at increasing (or decreasing) the acceptability of a controversial standpoint for the listener or reader, by putting forward a constellation of propositions intended to justify (or refute) the standpoint before a rational judge" (van Eemeren et al., 1996, p. 5). Many authors have argued that argumentation plays an important role in the learning process (e.g., Iordanou et al., 2019; Schwarz, 2009), notably in science learning (Driver et al, 2000; Kuhn, 1993; Tippett, 2009).

Argumentation is thought to improve learning through cognitive and social mechanisms (Schwarz, 2009). At the cognitive level, arguing involves making explicit one's views, reasoning and assumptions. Such a "self-explanation effect" (Chi et al., 1989) has been shown to foster knowledge acquisition (Bisra et al., 2018). At the social level, argumentation enables individuals to engage in, and resolve, sociocognitive conflicts (Asterhan, 2018). Sociocognitive conflicts designate situations that "arise when people hold different views or ideas about the same object," and that have "the potential to promote learning, cognitive development, and positive social relations" (Butera et al., 2019, p. 145). Argumentative exchanges involve being confronted with rebuttals of one's arguments, and developing rebuttals against arguments and views one disagrees with (Stein & Miller, 2012). Doing so enables arguers to resolve sociocognitive conflicts and update their views

(Leitão, 2000). More generally, arguing usually involves the aim of convincing an audience, which can be motivating and engaging (Schwarz, 2009).

Despite these strong theoretical rationales and a general acceptance that arguing fosters learning (e.g., Asterhan & Schwarz, 2016; Jonassen & Kim, 2010; Schwarz, 2009), there is no compelling empirical evidence to date that argumentation straightforwardly favours knowledge acquisition, as measured by domain-specific knowledge tests (Asterhan & Schwarz, 2016; Larrain et al., 2021; Wecker & Fischer, 2014). The lack of a measurable impact of argumentation activities on domain-specific knowledge was explained by a lack of conceptual clarity regarding the definition of learning (and notably, how one measures learning outcomes) and the processes whereby argumentation favours learning (Asterhan & Schwarz, 2016; Wecker & Fischer, 2014).

Finally, argumentation may be particularly relevant to learning essentially contested concepts, that is, "concepts the proper use of which inevitably involves endless disputes about their proper uses […]" (Gallie, 1956, p. 169). By coining this concept, Gallie wished to provide a rational foundation on which to constructively discuss complex concepts such as "democracy" or "work of art." Thus, in the context of classes focused on teaching essentially contested concepts, confronting different views and perspectives is not only beneficial to the learning process, but may also be seen as an integral and essential component of the content to be mastered.

Improving argumentation in MOOC forums

Even though there is sparse evidence that argumentation fosters knowledge acquisition, argumentation is known to foster cognitive elaboration (Stegmann et al., 2012) and cognitive engagement (Schwarz, 2009). Given these positive outcomes, MOOC participants may benefit from engaging in argumentative discussions with fellow learners. The only spaces where such interactions might occur are discussion forums, where the interactions have, however, been reported to be rare and of poor quality when they do occur (e.g., limited to sharing and comparing information, as opposed to more elaborate interactions such as resolving potential conflicts or synthesising knowledge, Tawfik et al., 2017).

Our goal is not to increase the quantity of contributions in MOOC discussion forums, but to improve their argumentative quality. To our knowledge, no studies have directly attempted to improve the quality of argumentation in MOOC discussion forums. The literature offers insights, however. Many interventions aimed at improving the quality of argumentation have adopted a computer-supported approach (Noroozi et al., 2012; for meta-analyses, see Vogel et al., 2017; Wecker & Fischer, 2014). In this literature, a substantial amount of research focuses on asynchronous online exchanges (e.g., Choi et al.,

2005; de Wever et al., 2008; Jeong & Joung, 2007; Nussbaum et al., 2004; Stegmann et al., 2007; Suthers et al., 2008), which are key features of MOOC discussion forums.

Several strategies have been used to improve the quality of argumentation in asynchronous online discussions. For instance, before posting their contributions, learners have been asked to pick a message type (e.g., "clarification" or "counterargument," Choi et al., 2005) or note starter (e.g., starting their messages with "On the opposite side…", Nussbaum et al., 2004). Other strategies include providing interaction scripts (i.e., "instructional plans that specify and sequence collaborative learning activities", Stegmann et al., 2007, p. 422), or providing visual aids to build arguments (Jeong & Joung, 2007; Stegmann et al., 2007). Similarly, the teacher may attribute specific roles to learners (e.g., starter, summariser, moderator, theoretician and source searcher, de Wever et al., 2008; de Wever & Strijbos, 2021) in order to foster social construction of knowledge.

Most of these interventions, however, may be difficult to implement in MOOC discussion forums due to constraints posed by MOOCs. One primary limitation is that certain prominent MOOC hosting platforms, such as edX (the platform utilised in our intervention), do not permit researchers and teachers to modify the interface of the discussion forums. Consequently, the use of note starters (Nussbaum et al., 2004), choice of message types (Choi et al., 2005) and visual aids for argument construction (Jeong & Joung, 2007; Stegmann et al., 2007) is restricted. These interventions typically involve modifications to the message writing interface, such as having participants complete different boxes for their claims, justifications and qualifications (Stegmann et al., 2007), which is not feasible within the existing platforms' limitations.

Second, MOOC discussion forums are characterised by self-paced participation. In other words, there are no constraints as to when users attend the course (and *a fortiori* participate in the discussion forums). When a learner posts a message on a forum, other users might answer within hours, days or weeks. It is also possible that their message gets no answer at all. Similarly, the self-paced nature of MOOCs makes it difficult for researchers to coordinate students' participation in exercises (e.g., argumentation training). In sum, while past research on asynchronous online interactions exerted some control over the timing of the interactions (e.g., by having learners asynchronously interact during an 80-minute window, Stegmann et al., 2007), it seems difficult to exert such control in interventions intended for implementation in MOOCs.

Finally, MOOC audiences are more diverse than usual class audiences in terms of first language, level of education, geographical location and cultural background. This diversity also poses a challenge when it comes to fostering argumentative discussions. For instance, interventions aimed at attributing roles to foster constructive interactions were previously conducted on relatively homogeneous populations (e.g., freshman students in Instructional Science, de Wever et al., 2008). Arguably, such an approach might be less well adapted to

contexts characterised by a higher degree of learner diversity, particularly in terms of academic background knowledge.

Developing an intervention adapted to MOOC constraints

To address these constraints, we developed an intervention aimed at improving learners' *basic* argumentation skills in *individual* exercises *before* they develop arguments in discussion forums. Focusing on basic argumentation skills enabled us to circumvent the constraint of the population's diversity, as it would not require elaborate background knowledge. Designing individual argumentation exercises (rather than group exercises) enabled us to develop an intervention that would not require participants to take part in the intervention simultaneously. Finally, having learners take part in the intervention before they take part in the discussion forums means that the interface of the forums does not need to be modified.

On a theoretical level, our intervention relies on the simplified version of Toulmin's Model of Argument (Weinberger & Fischer, 2006). Despite its limitations (see Jonassen & Kim, 2010; Leitão, 2000), this model suited our purpose for several reasons. First, it has the advantage of providing a framework within which to analyse single arguments. It therefore appeared appropriate for analysing MOOC forum contributions, given that interactions between learners tend to be sparse (Galikyan et al., 2021). As a result, a framework that enabled us to analyse single contributions seemed suitable. Second, given its simplicity, Toulmin's simplified model enabled us to design exercises that do not require learning of complex concepts. By contrast, other approaches, such as Blair and Johnson's (1987) argumentative principles – acceptability, relevance and sufficiency – require notions of logic. Finally, the simplicity of this model makes it combinable with other approaches (e.g., approaches centred on argument sequences, e.g., Leitão, 2000; see Weinberger & Fischer, 2006). As such, we viewed this model as an ideal first step in the development of interventions incorporating both single-argument and argument-sequence approaches to argumentation in the MOOC context.

Overview of the research

To our knowledge, while many studies have attempted to improve the formal quality of arguments in asynchronous online settings, none have done so in the ecological context of MOOCs. In this research, we sought to improve the quality of arguments in the discussion forums of a MOOC introducing an essentially contested concept, namely Corporate Social Responsibility (Lindgreen & Swaen, 2010). There is no unique, consensual definition of what constitutes corporate social responsibility.

For this purpose, we developed and experimentally tested an intervention aimed at improving the quality of MOOC users' arguments. We sought to create an intervention

adapted to the specific constraints of the MOOC forum environment. The intervention consisted in an introduction to the three formal components of arguments – claims (i.e., the position one is defending), justifications (i.e., the reasons brought forward to substantiate one's claim) and qualifications (i.e., limitations of the domain of validity of one's claim) – in accordance with the simplified version of Toulmin's Model of Argument (Stegmann et al., 2007; Toulmin, 1958; Weinberger & Fischer, 2006). The intervention also included exercises focused on the epistemic quality of justifications and qualifications (i.e., what distinguishes sound justifications/qualifications from poor ones). This aspect is notably absent from Toulmin's formal approach to argumentation, as highlighted by Backman et al. (2023).

First, we expected to observe a learning benefit in a context closely related to the exercises – namely a *near transfer* (Perkins & Salomon, 1992). Specifically, we expected that straight after completing exercise sessions introducing the components of an argument, participants in the experimental group would be more likely to justify (H1) and qualify (H2) their claims in response to open-ended questions aimed at prompting argumentation. We also expected to observe a learning benefit in the discussion forums, or a *far transfer*, so that participants in the experimental group would be more likely to justify (H3) and qualify (H4) their claims in forum posts.

The intervention included exercises regarding the quality of justifications and qualifications. Hence, we also expected that participants in the experimental group would report more scientific references in their justifications (H5) and would be more likely to use qualifications that refer to precise (rather than vague) conditions of validity (H6).

This study follows the ethical principles outlined in the APA Code of Conduct and complies with the national ethics guidelines of Belgium. The study did not involve any physical or psychological discomfort for participants.

Method

Participants and data

Data collection took place during the second semester of the 2022-2023 academic year, between February 11th and July 1st. The intervention was mandatory for management master's students at the Université catholique de Louvain (Belgium), but also open to anyone enrolled in the MOOC and willing to participate. One hundred and ten French-speaking participants completed the first questionnaire of the experimental and the control groups. The majority were students pursuing a master's degree in management (n = 57). The rest were either attending the paid version of the class (n = 10) or the free version (n = 43). Participants were randomly allocated to the experimental group (n = 56, 33 students, 17 attending the free version, 3 attending the paid version) and the control

group (n = 54, 27 students, 19 attending the free version, 7 attending the paid version). For each status, allocation did not significantly differ across condition, ps > . 24.

Eighty-four participants remained at the end of the intervention (i.e., completed the third questionnaire of the experimental group and the second questionnaire of the control group). They were evenly distributed across the experimental (n = 42, 26.3% dropout) and control groups (n = 42, 22.2% dropout). In line with the self-paced nature of MOOCs, participants exhibited considerable variability in the time span between the start of the first exercise session and termination of the last exercise session (M = 9.90 days, SD = 17.6; *Median* = 2.14; min = .01, max = 84.2). With regard to the textual corpus, we analysed a total of 255 written arguments (28,538 words). In both groups, participants had to write two arguments as individual exercises before progressing to the discussion forums, resulting in a total of 171 arguments ($n_{1st\ argument}$ = 100; $n_{2nd\ argument}$ = 71). The contributions in the two discussion forums amounted to 84 contributions ($n_{1st\ forum}$ = 43; $n_{2nd\ forum}$ = 41).

Design and procedure

The intervention was implemented in the first chapter of the MOOC "Discovering Corporate Social Responsibility" (*Découvrir la Responsabilité Sociétale des Entreprises*). This chapter is entitled "Defining the concept of CSR" (*Définir le concept de RSE*). It consists of eight subchapters (numbered from 1.1. to 1.8.), each subdivided into several short videos (5 to 10 minutes) and texts introducing the various dimensions of the definition of CSR.

Participants were randomly allocated to two groups at the beginning of the MOOC (see Figure 1). In the experimental group, participants had to complete three argumentation training sessions located at the end of sections 1.2. (session 1 – introduction), 1.3. (session 2 – consolidation) and 1.5. (session 3 – consolidation and extension). At the beginning of each training session, the intervention was presented as aiming to improve participants' argumentation skills. The training sessions introduced participants to three formal components of arguments (claims, justifications, qualifications) using a combination of brief theory readings, multiple-choice exercises and detailed feedback for each exercise (see details below). At the end of the first and third training sessions, participants were asked to answer open-ended questions aimed at prompting argumentation (see Figure 1). In the control group, participants only had to answer to the open-ended questions. The intervention was completed before participants were introduced to the discussion forums (in subchapters 1.6. and 1.7.).

At the end of the chapter, participants were asked to take part in two discussion forums as part of the class. The forums were part of the MOOC before the research was conducted and were not altered when the intervention was implemented. In the first forum (at the end of subchapter 1.6.), participants were asked to write down their criticisms of CSR.[a] In the

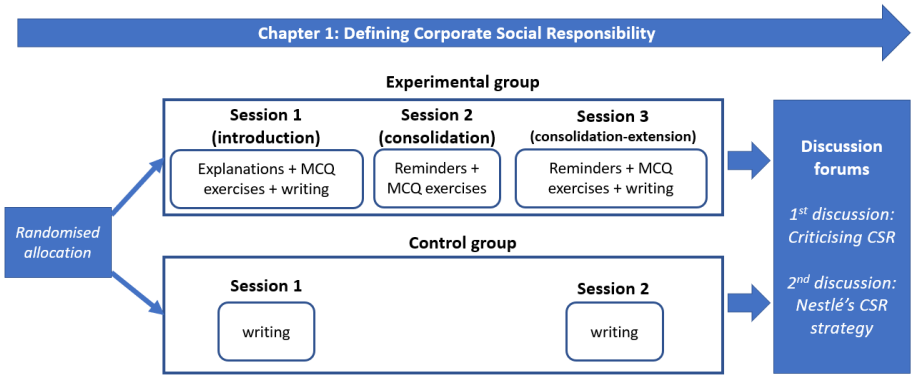

Fig. 1 Design of the intervention. MCQ stands for multiple choices questions. CSR stands for Corporate Social Responsibility

second forum (at the end of subchapter 1.7.), participants were asked to share their opinion (evaluation, criticisms) regarding Nestlé's CSR strategy.[b] Before taking part in the second forum, learners watched a video clip in which a representative of Nestlé described the company's CSR strategy.

The pacing of the participation differed between students and others. In the case of students pursuing the Master of Management, the professor in charge of the course encouraged them to take part in the online course in parallel with the weekly face-to-face classes. Students following this instruction would have completed the intervention within the first two weeks of February 2023. However, this pacing was not mandatory. The requirement for management students was to finish the five chapters of the course before the final evaluation in early May 2023. For participants who were not students, there was no particular instruction regarding the pacing of their participation.

Argumentation training sessions

Given the impossibility of providing individual feedback to learners, exercises consisted in multiple-choice questions. This format enabled us to use standardised feedback for correct or incorrect answers. The first author created the exercises, in close collaboration with the co-authors, as well as with the professor in charge of the MOOC and one of her PhD students. All exercise sessions were pretested and improved multiple times before the final version was implemented in the MOOC. The complete training sessions, along with associated feedback, can be found on the Open Science Framework: https://osf.io/89pqn/.

In the first exercise session (introduction session, see Figure 1), participants in the experimental condition were successively introduced to the notions of claim (in the first exercise), justification (in the second exercise) and qualification (in the third exercise). At the beginning of each exercise, participants read a brief description of the target formal

component and were shown a few examples related to the course content. For instance, in the first exercise – which introduced the notion of claim – participants first read the following:

> *The basis of an argument is a claim that one puts forward and that one wishes to defend, to establish as true. It must be stated as clearly as possible.*
>
> *Here are some examples:*
>
> *1) The primary objective of a company is to maximise shareholder value*
>
> *2) A company must above all think about its contribution to society*

Participants were then asked to identify the claim in an argument by selecting the appropriate number:

> *"[1] Our company cannot afford to have any priority other than profit. [2] We are in financial difficulty and other priorities could push us into bankruptcy."*
>
> *What is the claim?*

In the next exercise, they were first introduced to the notion of justification, then asked to identify the claim and the justification in an argument. In the third question, they were introduced to the notion of qualification and asked to identify all three components of an argument. After completing the three exercises, participants were shown a table that summarised the definitions and nature of all three components (see Figure 2).

The exercise session ended with an open-ended task where participants were asked to write a 500- to 1,000-character argument answering the following question: "What is the

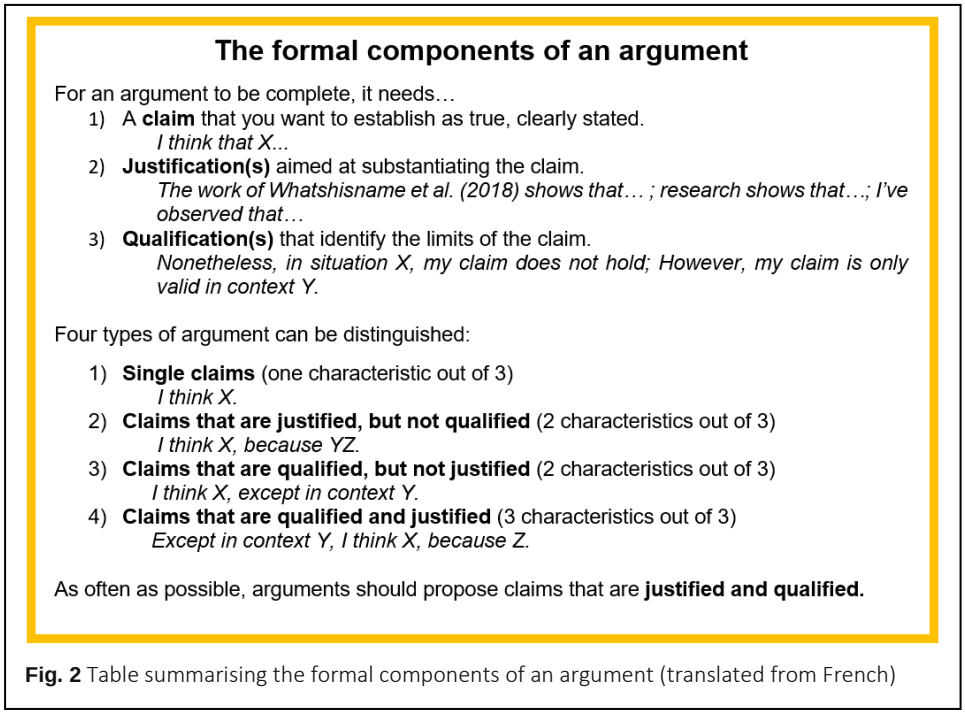

Fig. 2 Table summarising the formal components of an argument (translated from French)

purpose of a company?" Before answering, they were reminded of two classic positions regarding this issue: one prioritising shareholders' profit (advocated by neoclassical economist Milton Friedman), the other prioritising a positive impact on society (advocated by former Danone CEO Emmanuel Faber).

In the second exercise session (consolidation session), participants completed exercises aimed at strengthening their knowledge of the formal components of arguments. The exercises took a slightly different form from those in the first session: participants had to identify the missing component of arguments, as illustrated below:

> *"Argument: CSR is more than just compliance with laws and norms. Compliance with laws and conventions is a requirement for CSR, but a socially responsible company goes beyond compliance with laws and implements its own CSR initiatives."*
>
> *If this argument is incomplete, please identify the missing component(s).*

Participants identified the missing component (here, qualifications) via multiple-choice questions (the options being "justification(s)", "qualification(s)", and "none of the components are missing"). In addition, they were asked to characterise the arguments as either:

1) an isolated claim (with no justification or qualification)
2) a justified claim (not qualified)
3) a qualified claim (not justified)
4) a justified and qualified claim.

Since participants did not have access to the definition of the components in the instructions for the question as in the first sequence, they were given the opportunity to see the table summarising the different components of an argument (see Figure 2).

In the third and final exercise session (consolidation-extension session), participants had to identify the claim, justification and qualification of three arguments (two formally incomplete, one complete). These arguments took the same form as in the first session:

> *"Argument: [1] We cannot develop CSR if companies do not balance the interests of shareholders against those of other stakeholders. [2] Shareholder interests are often at odds with the interests of more vulnerable stakeholders.*
>
> *Please identify the segments corresponding to the different components of the argument: claim, justification, qualification."*

In addition, two exercises introduced participants to the issue of the quality of qualifications and justifications. Regarding justifications, participants read a brief text explaining the hierarchy of justifications, and notably, the superiority of scientific sources over personal opinions. Regarding qualifications, we emphasised the importance of identifying specific conditions under which the claim is not valid, rather than merely acknowledging the existence of exceptions (e.g., "Generally speaking, …"). In two exercises, participants had to rank three arguments based on the quality of their

justifications and qualifications. Here is the question focusing on the quality of justifications:

> *"Justifications can also vary in quality. Assessing the quality of justifications is a complex task. Here are a few simple guidelines.*
>
> *A justification based on true information will be superior to one based on false information. While it is often impossible to know directly whether information is true or false, the source of the information is an important indicator of reliability.*
>
> *On a scientific issue, information derived from clearly identified scientific studies […] is more likely to be true than a justification based on information found on an anonymous blog, or on your personal opinion.*
>
> *As much as possible, you should therefore cite the sources on which your justifications are based: this allows you to check the quality of the justification(s) put forward. […]"*

Participants were then asked to "rank the following arguments, from best to worst, based on the quality of their justification":

A) "According to Prof. François Maon, the priority criteria described in the Mitchell et al. (1997) model tend to overlook the interests of "weak" stakeholders. We therefore need to think about how to integrate the interests of these stakeholders into corporate priorities.

B) You only need to look at the world for five minutes to see that the interests of weak stakeholders are always ignored by companies. So we need to think about how to integrate the interests of these stakeholders into corporate priorities.

C) Traditional stakeholder approaches neglect the defence of "weak" stakeholders (such as local communities). We therefore need to think about how to integrate the interests of these stakeholders into corporate priorities."

The third exercise session closed with a second open-ended question asking participants to position themselves in relation to Mitchell et al.'s (1997) model of stakeholder identification and salience, which is a major theoretical model used in management.

In multiple-choice questions, we observed 18.11% of incorrect answers on average (i.e., all exercises included). Proportions of incorrect answers ranged from 4.76% to 42.86%. Exercises in which we asked participants to identify which component was missing from the argument (in training session 2) were much more difficult for participants than exercises in which we asked them to identify the different components of the argument. The first training session was clearly the easiest for participants, with only 8.84% of incorrect answers on average. In exercise session 2, we observed more incorrect answers (24.4%). In session 3, the rate of incorrect answers was 17.99%.

Coding the formal quality of arguments: analytic strategy

In order to assess the quality of participants' arguments in the discussion forums and individual exercises, we binary-coded the presence (vs absence) of claims, justifications and qualifications in arguments. Based on these three codes, we created a code indicating whether the argument was formally complete (i.e., justified and qualified). Two additional codes signalled the presence (vs absence) of scientific references in the justifications (i.e., high-quality justifications) and of clearly identified conditions of validity (i.e., high-quality qualifications). Those codes were decided a priori, before data analyses. In total, each argument was characterised on the basis of six binary-coded features (1 = present; 0 = absent):

1) Presence of a claim (What is the point that the arguer wishes to establish as true?)[c]
2) Presence of justification(s) (What are the reasons brought forward to substantiate the claim?)
3) Presence of qualification(s) (Does the author of the argument acknowledge limits to their claim?)
4) Complete argument (Is the claim justified and qualified?)
5) Presence of a high-quality justification (i.e., Does the author of the argument mobilise a scientific reference in their justification?)
6) Presence of clearly identified limitations to one's claim (Does the author of the argument identify conditions under which the claim would not be valid?)

A detailed description of how the formal components of arguments were identified is available in the online supplements. The coding was carried out by two independent judges. The coders were not aware of the participants' group (experimental or control) when coding the arguments. The coding was carried out separately by the two coders, who discussed their disagreements until they reached a consensus. Before discussion, the average Cohen's kappa was $\kappa = .63$ (*median* = .65, *min* = .27, *max* = .93).[d]

Results

Confirmatory analyses

Since the analyses only involve binary-coded variables, we tested our hypotheses using Chi-squared tests. When the conditions of applications were not met, we used Fisher's exact test.

Individual arguments

The results are summarised in Table 1, where we also included sensitivity analyses for the different arguments. For the first argument ("What is the purpose of a company?", $n_{argument} = 100$), participants in the experimental group were neither more likely to justify

Table 1 Effect of experimental condition on the prevalence of formal components in arguments

Effect size detected under a power of .80			Experimental		Control		
Exercise 1 (*n* = 100)			***n* = 51**	**%**	***n* = 49**	**%**	**χ^2**
The purpose of a company		Complete argument	28	54.9	20	40.86	1.99, *p* = .16
		justified	44	86.27	37	75.51	1.88, *p* = .17
	Phi = 0.28	Reference	17	33.33	7	14.29	**4.97, *p* = .026**
		qualified	30	58.82	27	55.1	0.14, *p* = .71
		Strong qualification	16	31.37	10	20.41	1.56, *p* = .21
Exercise 2 (*n* = 71)			***n* = 34**		***n* = 37**		
Self-positioning regarding the model of		complete arguments	28	82.35	30	81.08	0.13, *p* = .72
Mitchell et al. (1997)		justified	32	94.12	36	97.3	0.44, *p* = .51
	Phi = 0.33	Reference	2	5.88	0	0	*p* = .23*
		qualified	27	79.41	30	81.08	0.03, *p* = .86
		Strong qualification	7	20.59	6	16.22	0.288, *p* = .59
Forum 1 (*n* = 43)			***n* = 25**		***n* = 18**		**χ^2**
Criticisms of CSR		complete arguments	9	36	10	55.56	1.62, *p* = .20
		justified	19	76	14	77.78	0.02, *p* = .89
	Phi = 0.43	Reference	0	0	1	5.56	*p* = .39*
		qualified	11	44	11	61.11	1.23, *p* = .27
		Strong qualification	1	4	3	16.67	*p* = .28*
Forum 2 (*n* = 41)			***n* = 24**		***n* = 17**		
Opinion on Nestlé's CSR strategy		complete arguments	9	37.5	9	52.94	0.96, *p* = .33
		justified	15	62.5	15	88.24	3.36, *p* = .067
	Phi = 0.44	Reference	0	0	1	5.88	*p* = .44*
		qualified	17	70.83	12	70.59	0.00, *p* = .99
		Strong qualification	1	4.17	3	17.65	*p* = .30*

Note: Phi = .10, .30, .50 correspond to small, medium and large effect sizes, respectively. In bold is the significant difference observed in the results. Asterisks (*) indicate the use of Fisher's exact test due to insufficient number of expected observations in cells (< 5, McDonald, 2014).

their claims, $\chi^2(1) = 1.88$, $p = .17$ (86.3% justified arguments vs 75.5%) nor more likely to qualify their claims, $\chi^2(1) = 0.14$, $p = .71$ (58.8% qualified arguments vs 55.1%). Overall, they were not more likely to write formally complete arguments, $\chi^2(1) = 1.99$, $p = .16$ (54.9% complete arguments vs 40.9%). However, participants were more likely to refer to a scientific source to justify their claim, $\chi^2(1) = 4.97$, $p = .026$ (33.3% referring to a scientific source vs 14.3%). They were not more likely to use strong qualifications, $\chi^2(1) = 1.56$, $p = .21$ (31.4% referring to precise conditions of validity vs 20.4%).

In the second individual argument (self-positioning towards the management model of Mitchell et al., 1997, $n_{arguments} = 71$), participants in the experimental group were not more likely to justify their claims, $\chi^2(1) = 0.44$, $p = .51$ (94.1% of justified arguments vs 97.3%). They were not more likely to qualify their claims, $\chi^2(1) = 0.03$, $p = .86$ (79.4% of qualified arguments vs 81.1%). Overall, they were not more likely to write formally complete arguments, $\chi^2(1) = 0.24$, $p = .72$ (82.4% vs 81.1%). Finally, they were not more likely to use scientific references in their justifications, $p = .23$ (5.9% vs 0%), or strong qualifications, $\chi^2(1) = 0.29$, $p = .59$. (20.59% vs 16.22%).

In summary, completing exercises aimed at learning to identify the formal components of arguments did not lead to an improvement in the formal quality of arguments written directly after the exercise sessions.

Discussion forums

In the "criticisms of CSR" forum ($n_{arguments} = 43$), participants in the experimental group were not more likely to justify, $\chi^2(1) = 0.02$, $p = .89$ (76% of justified arguments vs 77.8%) or to qualify their claims, $\chi^2(1) = 1.23$, $p = .27$ (44% of qualified arguments vs 61.1%). Overall, they were not more likely to write formally complete arguments, $\chi^2(1) = 1.62$, $p = .20$ (36% of formally complete arguments vs 55.6%). As indicated by Fisher's exact tests (see Table 1), they were not more likely to mobilise scientific sources in their justifications, $p = .39$ (0% vs 5.56%), nor to use precise qualifications, $p = .28$ (4% vs 16.67%).

Similarly, in the "Nestlé's CSR strategy" forum ($n_{arguments} = 41$), participants in the experimental group were not more likely to justify, $\chi^2(1) = 3.36$, $p = .067$ (62.5% of justified arguments vs 88.2%) or qualify, $\chi^2(1) = 0.00$, $p = .99$ (70.8% of qualified arguments vs 70.6%), their claims. They were not more likely to write formally complete arguments, $\chi^2(1) = 0.96$, $p = .33$ (37.5% of formally complete arguments vs 52.9%). They were not more likely to mobilise scientific sources in their justifications, $p = .44$ (0% vs 5.9%) or to use strong qualifications, $p = .30$ (4.2% vs 17.7%).

Overall, participants in the experimental group did not demonstrate a formal improvement in their arguments compared to participants in the control group. We did not

observe the expected near learning transfer in the individual exercises, nor the far transfer in the discussion forums.

Exploratory analyses

Exercise sessions were longer for participants in the experimental group due to the argument components identification exercises. This may have affected the motivation of participants in the experimental group, and one could have anticipated a reduced time commitment to the exercises. However, the writing time did not significantly differ between conditions for the first individual argument, $t(108)^e = 0.45$, $p = .37$ ($M_{experimental} = 653$ seconds, SD = 469; $M_{Control} = 573$, $SD = 452$) or for the second individual argument, $t(81.3) = 0.45$, $p = .66$ ($M_{experimental} = 579$, SD = 504; $M_{Control} = 625$, $SD = 510$). Similarly, average word count did not differ between groups for the first individual argument, $t(99) = -0.04$, $p = .97$ ($M_{experimental} = 112$ words, $SD = 42.5$; $M_{Control} = 112$, $SD = 41.8$) or for the second individual argument, $t(66.3) = 1.87$, $p = .066$ ($M_{experimental} = 118$ words, $SD = 42.5$; $M_{Control} = 137$, $SD = 40.3$). Neither did it differ in the case of contributions in the first forum, $t(42.1) = -0.92$, $p = .36$ ($M_{experimental} = 91.4$ words, $SD = 48.1$; $M_{Control} = 78.7$, $SD = 44.2$) and the second forum $t(39.2) = 0.017$, $p = .987$ ($M_{experimental} = 91.2$ words, $SD = 63.4$; $M_{Control} = 94.8$, $SD = 32.4$). Thus, taking part in the exercise sessions before having to write arguments did not reduce the time spent writing or the length of participants' arguments throughout the intervention.

In an additional analysis on the two individual arguments ($n = 172$), we found a significant effect of the question ("What is the purpose of a company?" vs "Position yourself regarding the model of Mitchell et al.") on the formal quality of arguments. In the second argument, participants were more likely to justify, $\chi^2(1) = 8.26$, $p = .004$ (95.8% justified arguments vs 81%) and qualify, $\chi^2(1) = 10.5$, $p = .001$ (80.5% qualified arguments vs 57%) their claims. As a result, the second argument was more likely to be formally complete, $\chi^2(1) = 19.3$, $p < .001$ (82.8% formally complete vs 48%). In the first individual argument, however, participants were more likely to refer to a scientific source, $\chi^2(1) = 14.7$, $p < .001$ (24% arguments mobilising a scientific source vs 2.8%).

Discussion

In this research, we sought to improve the quality of arguments on the discussion forums of MOOCs exploring essentially contested concepts. For this purpose, we developed an intervention aimed at improving the argumentation skills of people attending a MOOC on Corporate Social Responsibility (i.e., an essentially contested concept). The intervention was designed to take into account the specific constraints of the MOOC setting, namely its fixed forum interface, its self-paced nature and its diversity of users.

In the proposed intervention, we expected participants to learn about the components of arguments by reading their definitions and functions, and by identifying these components in various pedagogical examples grounded in the course content. After completing individual exercises, participants were asked to write a short argument on a topic related to the MOOC. For participants in the experimental condition, the writing task was aimed at enabling them to put into practice the newly acquired (or, for those who already knew about the formal components of arguments, refreshed) argumentation knowledge.

Our intervention did not yield the expected results, as we did not observe an effect of our intervention on the quality of arguments written directly after completing the exercise sessions or later in discussion forums. In other words, in our study, learning about the formal components of arguments through a combination of readings and multiple-choice exercises with standardised feedback did not improve the quality of written arguments.

There are many potential reasons for the absence of significant results. A first potential explanation is the loss of statistical power due to many MOOC users not participating in the forums. This is in line with previous research showing that MOOC users seldom participate in discussion forums (Li et al., 2020; Tawfik et al., 2017). While the MOOC in which our intervention was implemented explicitly invited learners to participate in the forums, it still appeared that only half of the participants did so. Most of them were students attending the MOOC as part of their curriculum. Even in the hypothetical scenario where our intervention had an effect, the relatively small number of arguments in discussion forums reduced the probability of detecting it. The sensitivity analyses conducted revealed that under a power of 80% the sample sizes achieved only enabled us to detect medium (for individual exercises) or medium-to-large effects (for forum contributions). Hence, if our intervention had a small effect, it would be unlikely to be detected.

Another aspect that may partly explain our results is that the time lapse between students' participation in the different exercise sequences varied substantially. Some participants did all the exercise sequences within a few hours, others within a week, and still others within several months. This was expected, given the self-paced nature of MOOCs. Such variability most certainly represents an important source of between-participants variance, making effects harder to detect.

An additional potential issue is that our exercises mostly focused on *identifying* components of arguments in various pedagogical examples. Even though these exercise sequences ended with learners being asked to write an argument, *actually producing* (i.e., writing) claims, justifications and qualifications was not the focal point of the intervention. Future interventions might benefit from focusing more on the production, rather than the analysis, of arguments. This would reduce transfer distance between what is learned through the intervention and what is asked of them in the exercises and discussion forums (Perkins & Salomon, 1992). In the context of MOOCs, however, there is no possibility of

providing individualised feedback to participants. This is a significant limitation if one is seeking to develop an intervention in which students actively produce arguments.

Comparisons between arguments yielded some interesting exploratory findings. Participants were more likely to qualify and justify their claims in the second individual argument than in all the other questions. Hence, it seems as if the formal quality of arguments also depends on the prompt being used. The questions designed to prompt argumentation may have varied in the tasks they generated. In the first individual argument, participants were asked to position themselves between two conflicting positions regarding the purpose of a company. The task thus involved "taking sides". In the second individual argument, the task was to critically assess an influential theoretical model. In participants' responses, the qualifications of the argument typically consisted in an acknowledgement of the qualities of the model. These qualities were stressed during the online course. Asking learners to criticise a theory whose usefulness and qualities were stressed during the course may thus foster the qualification of arguments. Future research may examine how different question formats, relating differently to the content of the course, are susceptible to generating different tasks and, as a result, different argument formats. Rivera-Villa et al. (2024) have already shown that modifying forum instructions could foster cognitive engagement in forum contributions.

It is worth noting that overall, participants seldom included scientific references in the justifications of their arguments. An explanation for this result is the lack of access to such references during the course and during the exercises. The fact that participants were more likely to refer to a scientific reference in the first individual exercise can be explained by the fact that the references found in the arguments were also found in the examples used in those exercises (e.g., the IPCC report). So a way to improve argument quality might be to give students access to scientific references susceptible to being used in their arguments (e.g., by making them part of the course content or by providing students with a list of relevant references).

These findings emphasise that the context impacts the formal quality of arguments. In this regard, it might be interesting to distinguish between approaches that view argumentation as a set of skills that have to be trained (e.g., Kuhn, 1993) and approaches that view argumentation through its context of emergence (Berland & Hammer, 2012). The latter approaches assume that learners know how to argue (e.g., Sperber & Mercier, 2017), but that they do not do so in every context (Berland & Hammer, 2012). In this approach, the primary goal is not to teach learners argumentation skills, *but to create a context in which learners are willing to mobilise their pre-existing argumentation skills*. In the light of our results and past research (see Asterhan & Schwarz, 2016; Jonassen & Kim, 2010), an efficient way to foster argumentation on discussion forums would be to ask questions likely to elicit argumentation in the first place. For instance, rather than asking vague

questions (e.g., "What are your criticisms of CSR?"), it would be interesting to ask learners more specific and engaging questions (e.g., asking them to comment or position themselves in relation to a classic – and controversial – management theory). In this regard, research investigating how to foster cognitive engagement in MOOC discussion forums through forum instructions may be particularly relevant in the design of future interventions (e.g., Rivera-Villa et al., 2024).

Limitations and future directions

The choice of Toulmin's Model of Argument (Toulmin, 1958), i.e., an approach centred on individual arguments, was justified by the limited number of interactions observed in the MOOC discussion forums. However, future efforts to improve the quality of argumentation on discussion forums may benefit from combining approaches centred on single arguments and approaches centred on the construction of argumentative discussions (see Stegmann et al., 2007, for an example of such a combination of approaches). Relatedly, future research may benefit from training participants to evaluate the quality of arguments through different, complementary perspectives. For instance, participants could be trained to assess the acceptability of premises, or the relevance and sufficiency of the relationship between premises and conclusions (Blair & Johnson, 1987). On a similar note, complementing quantitative analyses of argument quality with qualitative examination of the content of arguments would be a valuable contribution.

Secondly, our assessment of the formal quality of arguments relied solely on binary-coded evaluation criteria (presence vs absence). This choice was motivated by the need to achieve acceptable inter-rater reliability. However, the trade-off for this reliability is a relatively basic evaluation of argument quality. Future research may refine ways of measuring argument quality, for instance by evaluating it on continuous, rather than dichotomous, variables.

Finally, it seems that participants had difficulty identifying which components were missing from an argument. This exercise format (unlike exercises in which participants were asked to identify the different components of an argument) was associated with a particularly high rate of incorrect answers. This question format may need to be either avoided or refined in future research.

Conclusion

While many studies have attempted to improve the quality of argumentation in asynchronous online settings (e.g., Choi et al., 2005; de Wever et al., 2008; Jeong & Joung, 2007; Nussbaum et al., 2004; Stegmann et al., 2007; Suthers et al., 2008), none – to our knowledge – have attempted to do so in the specific context of MOOC discussion forums. The absence of significant results and the limitations notwithstanding, this study reports

the first randomised intervention aimed at improving the quality of arguments in this context. Given the growing popularity of MOOCs (Voudoukis & Pagiatakis, 2022) and the benefits of engaging in argumentative discussions about course content (Schwarz, 2009), it is a scientific endeavour worthy of being pursued. Our research has helped to identify the specific challenges posed by MOOCs when it comes to developing and implementing interventions. Because of these constraints, many of the existing tools designed to improve the quality of students' argumentation in online asynchronous settings may be difficult to implement within the context of MOOC forums. We hope that this paper provides leads towards developing alternative ways to foster argumentation, by highlighting what may – or may not – be fruitful in this endeavour.

It is worth emphasising that despite these difficulties, MOOCs offer many opportunities for experimental research in education. Notably, they make it easy to implement interventions in pre-existing courses and to randomly allocate participants to different groups. Thus, in some respects, MOOCs are an ideal playground for experimental educational research. As such, it is worth striving to overcome the difficulties highlighted in our research.

Abbreviations

MOOC: Massively Open Online Class; CSR: Corporate Social Responsibility.

Endnotes

[a] "Share your criticisms of the concept and/or the way it is implemented in companies in the forum below. You can use real-life examples to illustrate your criticism, and do some research to back up your criticism, but avoid cutting and pasting."

[b] "What do you think of Nestlé's shared value creation strategy? Feel free to share your feelings and opinions about Nestlé's corporate social responsibility. What surprises you? What doubts or criticisms do you have? Don't forget to indicate in the title of your post whether your opinion is positive or negative, to make it easier for other learners to navigate and read."

[c] All the arguments analysed contained a claim. Therefore, we did not carry out analyses on this formal component of argument.

[d] The lowest reliability coefficient (κ = .27) corresponds to the presence of justifications in the second individual argument. For this exercise, participants often started by acknowledging the qualities of the model before introducing their criticisms. Criticisms often started with an adverb used to introduce qualifications in the first exercise (e.g., "however", "nevertheless", …). This led one of the raters to interpret the limits of the model as the claim, and the criticisms of the claim as the qualification (which encompassed justifications for participants' criticisms of the model). After discussion, raters agreed that the acknowledgement of the limits of the model constituted the qualifications of the argument, because participants were mostly focusing on identifying the limits of the model, and justifying their criticisms.

[e] Following the recommendation of Delacre et al. (2022), we carried out Welch's t-test instead of Student's.

Acknowledgements

We wish to warmly thank Pr. Valérie Swaen (for enabling us to implement our intervention in her MOOC), Pauline de Montpellier (for her feedback on the exercises), Dennis Rivera-Villa (for his feedback on the exercises), and Virginie Renson (for her technical help in the implementation of the intervention in the MOOC).

Authors' contributions

Kenzo Nera took part in the design of the intervention, devised the exercises and wrote the various versions of the manuscript. Mariane Frenay and Magali Paquot also took part in the design of the intervention and provided feedback on the exercises. They also provided feedback and direction on the successive versions of the manuscript.

Authors' information
Kenzo Nera is a FNRS postdoctoral researcher at the Université libre de Bruxelles (Belgium), in the CeSCuP (Center for Social and Cultural Psychology).

Mariane Frenay is a professor of educational psychology at the Université catholique de Louvain (Belgium), in the GIRSEF (Groupe interdisciplinaire de Recherche sur la Socialisation, l'Education et la Formation).

Magali Paquot is a permanent FNRS research associate in linguistics at the Université catholique de Louvain (Belgium). She is co-director of the CECL (Centre for English Corpus Linguistics).

Funding
Not applicable

Availability of data and materials
The data and materials used in this research are openly available at this address: https://osf.io/89pqn/.

Declarations

Competing interests
The authors declare that they have no competing interests.

Author details
[1] Psychological Sciences Research Institute, Université catholique de Louvain, Belgium

[2] Institute for Language and Communication, Université catholique de Louvain, Belgium

[3] Fonds de la Recherche Scientifique (FRS-FNRS), Belgium

Received: 16 October 2023 Accepted: 7 June 2024
Published online: 1 January 2025 (Online First: 15 July 2024)

References

Asterhan, C. S. (2018). Exploring enablers and inhibitors of productive peer argumentation: The role of individual achievement goals and of gender. *Contemporary Educational Psychology*, *54*, 66–78. https://doi.org/10.1016/j.cedpsych.2018.05.002

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

Backman, Y., Reznitskaya, A., Gardelli, V., & Wilkinson, I. A. (2023). Beyond structure: Using the rational force model to assess argumentative writing. *Written Communication*, *40*(2), 555–585. https://doi.org/10.1177/07410883221148664

Berland, L. K., & Hammer, D. (2012). Framing for scientific argumentation. *Journal of Research in Science Teaching*, *49*(1), 68–94. https://doi.org/10.1002/tea.20446

Bisra, K., Liu, Q., Nesbit, J. C., Salimi, F., & Winne, P. H. (2018). Inducing self-explanation: A meta-analysis. *Educational Psychology Review*, *30,* 703–725. https://doi.org/10.1007/s10648-018-9434-x

Blair, J. A., & Johnson, R. H. (1987). Argumentation as dialectical. *Argumentation*, *1*, 41–56. https://doi.org/10.1007/BF00127118

Butera, F., Sommet, N., & Darnon, C. (2019). Sociocognitive conflict regulation: How to make sense of diverging ideas. *Current Directions in Psychological Science*, *28*(2), 145–151. https://doi.org/10.1177/0963721418813986

Chi, M. T., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science*, *13*(2), 145–182. https://doi.org/10.1207/s15516709cog1302_1

Choi, I., Land, S. M., & Turgeon, A. J. (2005). Scaffolding peer-questioning strategies to facilitate metacognition during online small group discussion. *Instructional Science*, *33*, 483–511. https://doi.org/10.1007/s11251-005-1277-4

Collier, D., Daniel Hidalgo, F., & Olivia Maciuceanu, A. (2006). Essentially contested concepts: Debates and applications. *Journal of Political Ideologies*, *11*(3), 211–246. https://doi.org/10.1080/13569310600923782

Daniel, J. (2012). Making sense of MOOCs: Musings in a maze of myth, paradox and possibility. *Journal of Interactive Media in Education*, *2012*(3), 18. https://doi.org/10.5334/2012-18

Delacre, M., Lakens, D., & Leys, C. (2022). Correction: Why psychologists should by default use Welch's *t*-test instead of student's *t*-test. *International Review of Social Psychology*, *35*(1), 21. https://doi.org/10.5334/irsp.661

de Wever, B., & Strijbos, J. W. (2021). Roles for structuring groups for collaboration. In U. Cress, C. Rosé, A. F. Wise & J. Oshima (Eds.), *International handbook of computer-supported collaborative learning* (pp. 315–331). Springer. https://doi.org/10.1007/978-3-030-65291-3_17

de Wever, B., Schellens, T., Van Keer, H., & Valcke, M. (2008). Structuring asynchronous discussion groups by introducing roles: Do students act in line with assigned roles? *Small Group Research*, *39*(6), 770–794. https://doi.org/10.1177/1046496408323227

Driver, R., Newton, P., & Osborne, J. (2000). Establishing the norms of scientific argumentation in classrooms. *Science Education*, *84*, 287–312. https://doi.org/10.1002/(SICI)1098-237X(200005)84:3<287::AID-SCE1>3.0.CO;2-A

Galikyan, I., Admiraal, W., & Kester, L. (2021). MOOC discussion forums: The interplay of the cognitive and the social. *Computers & Education*, *165*, 104133. https://doi.org/10.1016/j.compedu.2021.104133

Gallie, W. B. (1956). Essentially contested concepts. *Proceedings of the Aristotelian Society*, *56*, 167–198. https://www.jstor.org/stable/4544562

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

Jeong, A., & Joung, S. (2007). Scaffolding collaborative argumentation in asynchronous discussions with message constraints and message labels. *Computers & Education*, *48*(3), 427–445. https://doi.org/10.1016/j.compedu.2005.02.002

Jonassen, D. H., & Kim, B. (2010). Arguing to learn and learning to argue: Design justifications and guidelines. *Educational Technology Research and Development*, *58*, 439–457. https://doi.org/10.1007/s11423-009-9143-8

Kuhn, D. (1993). Science as argument: Implications for teaching and learning scientific thinking. *Science Education*, *77*(3), 319–337. https://doi.org/10.1002/sce.3730770306

Larrain, A., Singer, V., Strasser, K., Howe, C., López, P., Pinochet, J., Moran, C., Sánchez, Á., Silva, M., & Villavicencio, C. (2021). Argumentation skills mediate the effect of peer argumentation on content knowledge in middle-school students. *Journal of Educational Psychology*, *113*(4), 736–753. https://doi.org/10.1037/edu0000619

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

Li, H., Kim, M. K., & Xiong, Y. (2020). Individual learning vs. interactive learning: A cognitive diagnostic analysis of MOOC students' learning behaviors. *American Journal of Distance Education*, *34*(2), 121–136. https://doi.org/10.1080/08923647.2019.1697027

Lindgreen, A., & Swaen, V. (2010). Corporate social responsibility. *International Journal of Management Reviews*, *12*(1), 1–7. https://doi.org/10.1111/j.1468-2370.2009.00277.x

López Meneses, E., Vázquez Cano, E., & Mac Fadden, I. (2020). MOOC in higher education from the students' perspective. A sustainable model? In J. L. S. Sánchez-Serrano, F. Maturo & S. Hošková-Mayerová (Eds.), *Qualitative and quantitative models in socio-economic systems and social work* (pp. 207–223). Springer. https://doi.org/10.1007/978-3-030-18593-0_17

McDonald, J. H. (2014). *Handbook of biological statistics*. Sparky House.

Mitchell, R. K., Agle, B. R., & Wood, D. J. (1997). Toward a theory of stakeholder identification and salience: Defining the principle of who and what really counts. *Academy of Management Review*, *22*(4), 853–886. https://doi.org/10.5465/amr.1997.9711022105

Noroozi, O., Weinberger, A., Biemans, H. J., Mulder, M., & Chizari, M. (2012). Argumentation-Based Computer Supported Collaborative Learning (ABCSCL): A synthesis of 15 years of research. *Educational Research Review*, *7*(2), 79–106. https://doi.org/10.1016/j.edurev.2011.11.006

Nussbaum, E. M., Hartley, K., Sinatra, G. M., Reynolds, R. E., & Bendixen, L. D. (2004). Personality interactions and scaffolding in on-line discussions. *Journal of Educational Computing Research*, *30*(1-2), 113–137. https://doi.org/10.2190/H8P4-QJUF-JXME-6JD8

Perkins, D. N., & Salomon, G. (1992). Transfer of learning. In T. Husén & T. N. Postlethwaite (Eds.), *The international encyclopedia of education* (pp. 425–441). Pergamon.

Rivera-Villa, D., Frenay, M., & Swaen, V. (2024). The learning design of MOOC discussion forums: An analysis of forum instructions and their role in supporting the social construction of knowledge. *Technology, Knowledge and Learning*, *29*, 585–615. https://doi.org/10.1007/s10758-023-09670-w

Schwarz, B. B. (2009). Argumentation and learning. In N. M. Mirza & A. N. Perret-Clermont (Eds.), *Argumentation and education: Theoretical foundations and practices* (pp. 91–126). Springer. https://doi.org/10.1007/978-0-387-98125-3_4

Sperber, D., & Mercier, H. (2017). *The enigma of reason*. *A new theory of human understanding*. Harvard University Press.

Stegmann, K., Wecker, C., Weinberger, A., & Fischer, F. (2012). Collaborative argumentation and cognitive elaboration in a computer-supported collaborative learning environment. *Instructional Science*, *40*, 297–323. https://doi.org/10.1007/s11251-011-9174-5

Stegmann, K., Weinberger, A., & Fischer, F. (2007). Facilitating argumentative knowledge construction with computer-supported collaboration scripts. *International Journal of Computer-Supported Collaborative Learning*, *2*, 421–447. https://doi.org/10.1007/s11412-007-9028-y

Stein, N. L., & Miller, C. A. (2012). I win–you lose: The development of argumentative thinking. In J. F. Voss, D. N. Perkins & J. W. Segal (Eds.), *Informal reasoning and education* (pp. 265–290). Routledge.

Suthers, D. D., Vatrapu, R., Medina, R., Joseph, S., & Dwyer, N. (2008). Beyond threaded discussion: Representational guidance in asynchronous collaborative learning environments. *Computers & Education*, *50*(4), 1103–1127. https://doi.org/10.1016/j.compedu.2006.10.007

Tawfik, A. A., Reeves, T. D., Stich, A. E., Gill, A., Hong, C., McDade, J., Pillutla, V. S., Zhou, X., & Giabbanelli, P. J. (2017). The nature and level of learner–learner interaction in a chemistry massive open online course (MOOC). *Journal of Computing in Higher Education*, *29*, 411–431. https://doi.org/10.1007/s12528-017-9135-3

Tippett, C. (2009). Argumentation: The language of science. *Journal of Elementary Science Education*, *21*(1), 17–25. https://doi.org/10.1007/BF03174713

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

van Eemeren, F. H., Grootendorst, R., Henkemans, F. S., Blair, J. A., Johnson, R. H., Krabbe, E. C. W., Plantin, C., Walton, D. N., Willard, C. A., et al. (1996). *Fundamentals of argumentation theory: A handbook of historical backgrounds and contemporary developments.* Lawrence Erlbaum Associates, Inc.

Vogel, F., Wecker, C., Kollar, I., & Fischer, F. (2017). Socio-cognitive scaffolding with computer-supported collaboration scripts: A meta-analysis. *Educational Psychology Review*, *29*, 477–511. https://doi.org/10.1007/s10648-016-9361-7

Voudoukis, N., & Pagiatakis, G. (2022). Massive open online courses (MOOCs): Practices, trends, and challenges for the higher education. *European Journal of Education and Pedagogy*, *3*(3), 288–295. https://doi.org/10.24018/ejedu.2022.3.3.365

Waldrop, M. M. (2013). The virtual lab. *Nature*, *499*(7458), 268–271. https://doi.org/10.1038/499268a

Wecker, C., & Fischer, F. (2014). Where is the evidence? A meta-analysis on the role of argumentation for the acquisition of domain-specific knowledge in computer-supported collaborative learning. *Computers & Education*, *75*, 218–228. https://doi.org/10.1016/j.compedu.2014.02.016

Weinberger, A., & Fischer, F. (2006). A framework to analyze argumentative knowledge construction in computer-supported collaborative learning. *Computers & Education*, *46*(1), 71–95. https://doi.org/10.1016/j.compedu.2005.04.003

Weinhardt, J. M., & Sitzmann, T. (2019). Revolutionizing training and education? Three questions regarding massive open online courses (MOOCs). *Human Resource Management Review*, *29*(2), 218–225. https://doi.org/10.1016/j.hrmr.2018.06.004

Publisher's Note