



Using the Controversy over Human Race to Introduce Students to the Identification and the Evaluation of Arguments

Pablo Antonio Archila¹ · Jorge Molina² · Giovanna Danies³ · Anne-Marie Truscott de Mejía⁴ · Silvia Restrepo¹

Accepted: 2 November 2021 / Published online: 13 November 2021
© The Author(s), under exclusive licence to Springer Nature B.V. 2021

Abstract

The identification and the evaluation of arguments are fundamental elements of critical thinking. However, the explicit promotion of these elements is virtually absent from university science courses. Much of the reason for this is that in most universities, across nearly all disciplines, instructors are required to see the conceptual content coverage of the syllabus as a priority. Moreover, lack of preparation and the fact that critical thinking activities are time-consuming rapidly reduce the interest of many instructors to include them in their courses. Here, we describe the use of a dialogue-based critical thinking classroom scenario (CTCS). The study used a mixed-methods approach with both quantitative and qualitative analyses of questionnaire responses. One hundred and seventeen undergraduates (73 females; 44 males; ages 16–24 years), enrolled in an introductory science course in Colombia, were asked to identify and evaluate arguments regarding a dialogue between two scientists who explore the controversial question of whether or not the concept of race is applicable to humans. It was found that the dialogue-based CTCS provided students with opportunities to identify and evaluate arguments both for and against the question and to make informed decisions about whether or not the concept of race in humans is biologically meaningful. Moreover, analyses of responses to closed-ended and open-ended questions revealed that more than half the participants were able to evaluate arguments in a fair-minded way. Practical implications for the cultivation of critical thinking skills in higher education and further research are discussed.

1 Introduction

In his book, *On Reasoning and Argument*, David Hitchcock (2017) presents us with the following statement:

An ideal “critical thinker” is open-minded and fair-minded, searches for evidence, tries to be well-informed, is attentive to others’ views and their reasons, propor-

✉ Pablo Antonio Archila
pa.archila@uniandes.edu.co

Extended author information available on the last page of the article

tions belief to the evidence, and is willing to consider alternatives and revise beliefs (p. 477).

This comment gives us an idea of how complex (and ambitious) the goal of preparing critical thinkers is. Within this complexity, argumentation is widely accepted as one among many elements that together contribute to this preparation. In other words, engaging students in argumentation practices (e.g., argument identification, argument evaluation) is a step in the right direction. However, there is still much work to be done and other elements that need to be explicitly worked on in order to truly prepare an ideal “critical thinker” equipped with the qualities highlighted by Hitchcock (2017). In this article, we report on a possible, effective, and pragmatic way to take this step, introducing students to the identification and the evaluation of arguments.

According to Andrews (2015), Archila (2018), Epstein (2017), Harker (2015), Hitchcock (2017), Jiménez-Aleixandre and Puig (2012), Lau (2011), and Walton (2019), the identification and the evaluation of arguments are fundamental elements among many other facets of critical thinking. These scholars assert that argumentation contributes to critical thinking and vice versa. To be clear, “argumentation implies criticality; the one cannot function without the other” (Andrews, 2015, p. 60). It is therefore not surprising to find that argumentation and critical thinking are recognized as key skills in well-known science education documents, such as *A Framework for K–12 Science Education: Practices, Crosscutting Concepts, and Core Ideas* (National Research Council, 2012) and *Next Generation Science Standards: For States, by States* (Next Generation Science Standards Lead States, 2013).

More recently, Archila et al., (2019, 2021a) and Hyytinen et al. (2019) claimed that critical thinking is essential for the development of scientific thinking skills. Thus, making informed decisions about climate change (Rehg, 2011), evolution and creationism (Archila & Molina, 2020), GMO crops (Fahnestock, 2020), nuclear energy (Jho et al., 2014), and genetics, ancestry, and race (Beckwith et al., 2017), among other controversies, requires educated citizens who critically identify and evaluate arguments as a key element of the construction of democratic societies. In this regard, university science courses can be considered as one of the multiple scenarios which involve students in authentic and meaningful educational practices aimed at providing them with explicit opportunities to identify and evaluate arguments.

The promotion of critical thinking is a complex and time-consuming process (Andrews, 2015; Hyytinen et al., 2019). This is one reason why some instructors prefer not to include critical thinking activities in their university science courses. Moreover, instructors are not prepared to promote higher-order cognitive-linguistic skills, such as argumentation and critical thinking (Archila, 2014a, 2014b). Another reason is that “the literature says surprisingly little about pedagogical principles of integrating critical thinking coherently in teaching and learning” (Hyytinen et al., 2019, p. 60). For the reasons just mentioned, it makes sense to provide research evidence on how to make critical thinking classroom scenarios (CTCSs) (Archila, 2018) an explicit practice of any university course interested in engaging students in the identification and the evaluation of arguments in a pragmatic and effective way. Thus, the present article outlines the effect of a dialogue-based CTCS in providing undergraduate students with opportunities to identify and evaluate arguments regarding a dialogue written by Beckwith et al. (2017), in which two scientists interact argumentatively to explore the controversial question of whether or not the concept of race in humans is biologically meaningful.

2 Conceptual Framework

In this section, we present the conceptual bases of the CTCS proposed in this study. These focused on six elements, namely: (1) argumentation, (2) critical thinking, (3) arguments identification, (4) arguments evaluation, (5) decision-making, and (6) controversial question. We start by clarifying the meaning of some terms which are widely used in our research. In the current article, we consider that “argumentation in scientific topics can be defined as the connection between claims and data through justifications or the evaluation of knowledge claims in light of evidence, either empirical or theoretical” (Erduran & Jiménez-Aleixandre, 2007, p. 13, as cited in Evagorou & Osborne, 2013). This definition is valuable as it differentiates argumentation from explanation which is defined “as providing an account of how or why a phenomenon occurs and explaining why the natural world works in particular ways” (McNeill, 2011, 795). With this in mind, it is important to clarify that argumentation about socioscientific issues (e.g., the controversy over human race) depends not only on mastery of scientific knowledge, but also on the consideration of moral and ethical values (Evagorou, 2011; Evagorou et al., 2012; Nielsen, 2012; Oliveira et al., 2012, as cited in Evagorou & Osborne, 2013). Moreover, we define an “argument” as the articulation of a claim with a piece of reason (reason-based argument) and/or evidence (evidence-based argument) (Erduran et al., 2020).

Archila (2018) defines a CTCS as “a set of teaching and learning conditions deliberately planned and implemented in a classroom to promote critical thinking” (p. 57). Additionally, and bearing in mind that Hitchcock (2017) reminds us that there is, as yet, no consensus on the definition of the notion of critical thinking and multiple perspectives is available, the following definition was adopted: “reasonable reflective thinking focused on deciding what to believe or do” (Ennis, 2015, p. 32).

2.1 Element One—Argumentation

In his iconic review published in *Science*, Jonathan Osborne (2010) gave strong reasons why the promotion of argumentation skills should be an imperative within science education. One reason is that argumentation is inherent to the practice of science. In this sense, he noted that “without argument and evaluation, the construction of reliable knowledge would be impossible” (p. 464). This largely explains why several scholars maintain that students should be provided with structured opportunities to engage in authentic argument evaluation practices in order to enrich their argumentation skills (e.g., Archila, 2015b; Greco Morasso, 2009; Jiménez-Aleixandre & Puig, 2012).

Another strong reason given by Osborne (2010) is that argumentation is a key ally of critical thinking. In line with this, Archila (2018) provides evidence for the claim that engaging students in argument evaluation is a concrete action to help them to think critically. This claim has been widely discussed by several scholars. Sampson and Schleigh (2013), for example, reiterate that instructors should encourage students to critique one another’s ideas and reflect with them about how to evaluate the arguments of others. They also invite us to explore and implement research-based, pragmatic, and effective activities to make the promotion of argument evaluation a reality in the science classroom rather than just rhetoric. In similar vein, Osborne (2012) argues that students become involved in higher-order cognitive processes, such as critical thinking and decision-making when they are provided with explicit opportunities to make judgments about the argument of others.

This is particularly important, since evidence suggest that not all the argumentation tasks are equally demanding for students (e.g., Berland and McNeill 2010; Sampson & Schleigh, 2013). Accordingly, Berland and McNeill (2010) contend that science instructors should be aware of the complexity behind an argumentation task in order to offer students better support.

In the case of argument evaluation, Sampson and Schleigh (2013) maintain that this is a complex task because it demands analyzing whether or not the evidence provided is relevant and effectively supports the claim. In this regard, Osborne (2012) points that critical thinking skills are cultivated when students are given opportunities to critique the arguments of others. Likewise, he insists that the development of students' argument assessment skills is a significant challenge because students tend to make use of biased argumentation. This typically occurs when they exclusively use arguments that support their points of views, ignoring possible counterarguments. In relation to this point, Evagorou and Osborne (2013) explain that socioscientific issues are open-ended issues without a clear-cut solution. This is why they are good candidates to engage students in productive argument evaluation, and thus promote a more critical view about these issues.

Additionally, Evagorou and Osborne (2013) note that engaging students in argumentation about socioscientific issues requires them to “see argumentation as the process of exploring ethical issues, a process which involves moral judgment about issues of scientific concern” (p. 212). Knight and McNeill (2015) argue that using socioscientific contexts to foster argument evaluation is beneficial for students to become aware of the moral, ethical, and political factors that influence these contexts. Most importantly, they maintain that the situation becomes even more complicated because these factors “frequently rely on non-empirical evidence” (p. 625). Therefore, the promotion of critical thinking skills (e.g., argument evaluation) is a fundamental condition to prepare students to deal with science-based social controversies in which there is no straightforward solution.

2.2 Element Two—Critical Thinking

Andrews (2015) argues that the promotion of critical thinking is an important but neglected goal of higher education. Much of the reason for this is that in theory, there is a broad agreement among higher education stakeholders that engaging students in critical thinking practices is vital for helping them to make informed and critical decisions about professional and personal issues. In practice, nonetheless, higher education courses provide students with very few opportunities to cultivate their critical thinking skills. In this regard, several scholars call for the promotion of critical thinking with concrete actions (Archila, 2018; Hitchcock, 2017; Jiménez-Aleixandre & Puig, 2012). For example, Archila (2018) provides research evidence for the claim that students cultivate their critical thinking skills when they are given opportunities to identify and evaluate arguments. At first glance, it seems irrelevant to engage students in argument identification. However, van Eemeren et al. (2015) have reported that identifying arguments can be a great challenge for some students. The importance of this is that, as Archila (2018) observes, a proper argument identification is fundamental for a desirable argument evaluation. In particular, argument evaluation has been largely accepted as a key element of critical thinking (Ennis, 2015; Jiménez-Aleixandre & Puig, 2012; Osborne, 2012). The reason is straightforward: argument evaluation demands judging whether or not the articulation of a claim with reason and/or evidence is coherent and valid within a process of argumentation. Hence, Andrews's (2015) idea that argumentation and critical thinking are intimately related makes sense.

Thus, an important question to ask is how to introduce the identification and the evaluation of arguments as a concrete action to cultivate critical thinking skills in higher education.

To answer this question, Hitchcock (2017) invites us to consider that there are two ways of developing critical thinking in an educational program. These are the two pure models proposed by Swartz and Perkins (1990). The first model “is infusion, where the strategies, skills, dispositions and attitudes of a critical thinker are developed in the context of subject-matter instruction” (Hitchcock, 2017, p. 489). He emphasizes that asking students to identify assumptions implicit in the reasoning of key argumentative texts from the area of a university course (e.g., economy, history, physics) might be an example of a characteristic activity of the infusion model. A clear advantage of this model is that it is pragmatic (realistic). To be precise, it creates a possibility for the instructor to cultivate students’ critical thinking without sacrificing the conceptual content coverage of the syllabus. Yet, a notable disadvantage of the infusion model is that it requires students’ knowledge of subject matter.

The second pure model is “stand-alone instruction,” which involves a separate course dedicated entirely to promoting critical thinking. Due to the nature of this model, the instructor is more likely to have time for planning specific activities centered on the multiple difficulties of the students to think critically in everyday life. This is an undeniable advantage. In undergraduate education, however, given that the professional context of a physician may be a little different to that of a historian, one problem of the stand-alone instruction model is that the idea of cultivating critical thinking skills in a separate course might exacerbate the obstacles students find in becoming aware of the importance of acknowledging critical thinking ability as a vital aspect in the professional contexts in which they will be involved after graduating.

Hitchcock (2017) concluded that a combination of infusion and separate instruction would seem ideal. He stresses that such a combination requires not only a strong commitment from educational institutions, but also from its senior academic leadership, to teaching critical thinking across the curriculum. It may be obvious to point out that this is hard to achieve. Nevertheless, higher education institutions can start by adopting one of the two models discussed. In our case, we decided to use the infusion model in our CTCS. Much of the reason for this is that we were particularly interested in providing research evidence for the claim that students can be introduced to the identification and the evaluation of arguments as a concrete action to cultivate their critical thinking skills without sacrificing the conceptual content coverage of the syllabus.

2.3 Element Three—Argument Identification

Having opted for the infusion model, it should be pointed out that several academics (e.g., Andrews, 2015; Archila, 2018; Epstein, 2017; Hitchcock, 2017; Jiménez-Aleixandre & Puig, 2012) agree that the identification and the evaluation of arguments to engage students in critical thinking practices are an under-researched possibility in higher education. Other important features to be considered in educational practices interested in providing students with opportunities to think critically in university courses include the following: critical spirit, fair-mindedness, open-mindedness, respect for alternative viewpoints, seeing both sides of an issue, taking a position and changing a position when the evidence and reasons are sufficient, and willingness to seek or be guided by reason (Davies & Barnett, 2015; Ennis, 2015; Hitchcock, 2017). These features offer an idea of the complexities behind the preparation of critical thinkers. Importantly, Andrews (2015) and Wendland et al. (2015) note that the idea of promoting critical thinking in university courses is only

now emerging, and therefore there is still much work to be done. For this reason, we limited our study to two aspects: (1) identification and (2) assessment of arguments.

According to Ennis (2015), ideal critical thinkers have the ability to analyze arguments. Such analysis requires the identification and the evaluation of arguments. Specifically, the identification process assists students in better understanding others' points of view. As part of a study conducted by van Eemeren et al. (2015), 40 s form pupils and 82 third form pupils in a lower stream in three comprehensive schools near Amsterdam were presented with 40 text fragments which consisted of one simple argumentation or a text of equal length and complexity. Each text is reproduced in the form of one compound sentence, divided by commas. For example: "In my opinion the presence of trees along the road is important, they reduce the tedium" (p. 736). Participants had to indicate in each text whether argumentation was present or not, and to underline (identify) the argument if present. One of the conclusions of the study was that identifying arguments was a real challenge for some students. Given this conclusion, it is probably not so surprising that van Eemeren et al. (2015) recommend the identification of arguments as a classroom activity.

2.4 Element Four—Argument Evaluation

In his book, *Argument Evaluation and Evidence*, Walton (2016) highlights the key role of evidence to assess arguments used in personal decision-making, law, scientific inquiry, and public debate. In the case of science education, Archila (2015a) suggests that argument evaluation is closely linked to evidence evaluation, with both requiring similar cognitive processes. Additionally, in a study conducted in France, Archila (2015b) showed results that supported the claim that evidence assessment promotes critical thinking. Most importantly, he found that the use of controversies in the classroom helped to raise the students' awareness of the essential role of evidence evaluation in the advancement of science.

In the current article, the expression "argument evaluation" is used to refer to "the verification of the logical validity and of the persuasiveness of arguments" (Greco Morasso, 2009, pp. 233–234). Argument assessment is a complex activity because it demands analyzing whether or not the articulation of a claim with a piece of the evidence is solid, strong, rational, and reasonable within an argumentation process. Relevance of intention (Jørgensen, 2007), formal representation (Selinger, 2014), artificial intelligence (Walton, 2015, 2016), and logical evaluation (Botting, 2016) are some of the perspectives from which such complexity has been discussed.

2.5 Element Five—Decision-Making

Recently, Archila (2018) has demonstrated that decision-making is a good starting point for the identification and the evaluation of the arguments of others. He mentions that there are at least three concrete approaches to involve students in decision-making whose use will depend on the objectives of the science teaching and learning session: (1) Each student makes a decision (e.g., Sakschewski et al., 2014), (2) each student makes a decision, then a small-group (or a whole-class) decision must be made (e.g., Jho et al., 2014), and (3) each student makes a decision before and after a small-group (or a whole-class) discussion (e.g., Archila, 2017). In our CTCS, each student makes a decision at the beginning and at the end of the activity without requiring a small-group or a whole-class discussion. This is why we decided to use the first approach in our CTCS (e.g., Sakschewski et al., 2014).

2.6 Element Six—Controversial Question

In a CTCS, a controversial question is a fundamental element. Archila (2018) maintains that this element is what effectively gives meaning to the actions of identifying and evaluating arguments and making more informed decisions. By a controversy, we mean an issue that will lead to a high level of different understandings among significant numbers of people (Harker, 2015). Archila (2015a) highlights the importance of presenting the students with a controversy in the form of a (1) provoking and (2) ambiguous question. The reason for this is that these two characteristics are relevant to help students focus on the pieces of evidence they use rather than the decisions they make about the controversial question. In the case of our CTCS, it revolves around the controversy over human race. More specifically, our scenario explores the controversial question of whether or not the concept of race in humans is biologically meaningful. In “the dialogue used in this project” section, we will explain what this controversy is about, why this was chosen, and how it was articulated with the six elements of our CTCS, namely: (1) argumentation, (2) critical thinking, (3) arguments identification, (4) arguments evaluation, (5) decision-making, and (6) controversial question.

3 Aims and Significance of the Study

The focus of the empirical study reported in the current article is framed in terms of the realistic and moderate purpose of this study. First, it aims to describe participants’ decisions about the controversial question of whether or not the concept of race in humans is biologically meaningful. Second, it aims to provide participants with opportunities to identify the arguments regarding a dialogue written by Beckwith et al. (2017), in which two scientists interact argumentatively to explore this controversial question. Third, it aims to engage students in the evaluation of the arguments presented in the dialogue. These aims emerged from the need for evidence to support the hypothesis that the evaluation of arguments from a dialogue about the controversy over the concept of race in humans could be a useful classroom scenario in taking one step, among many others that should be taken, towards the preparation of critical thinkers.

The significance of the present study is that it expands on the scope of some notable work carried out previously that has focused on the introduction of students to the identification and the evaluation of arguments. As such, the research questions that guided this exploratory investigation were:

- (1) Does the CTCS provide participants with opportunities to make a decision about the controversial question of whether or not the concept of race in humans is biologically meaningful?
- (2) Does the CTCS provide participants with opportunities to identify the arguments presented in a dialogue in which two scientists interact argumentatively to explore the controversial question?
- (3) Does the CTCS provide participants with opportunities to evaluate the arguments presented in a dialogue in which two scientists interact argumentatively to explore the controversial question?

4 The Dialogue Used in This Project

As mentioned previously, our CTCS revolves around the controversy over the concept of race in humans. Part of the controversy stems from the term “race” itself. For example, Barnshaw (2008) asserts that “a *race* is a social grouping of people who have similar physical or social characteristics that are generally considered by society as forming a distinct group” (p. 1091, italics in original). Marks (2010) provides us with another view mentioning that “race is a sense-making system imposed upon the facts of difference. Races are not merely human divisions, they are politically salient human divisions” (p. 271). In line with Marks’ (2010) view, Morning (2011) contends that “individuals do not carry race within them; instead, race is a label that is imposed on them (or a container into which they are put) depending on the society in which they find themselves” (p. 18).

Finally, Templeton (2016) observes that biologically, race is “a subpopulation within a species, also called a subspecies, that has sharp geographic boundaries separating it from the remainder of the species, with the boundaries characterized by a high degree of genetic differentiation defined either through a quantitative threshold or qualitatively as a separate evolutionary lineage” (p. 347). Importantly, Relethford (2013) highlights the fact that in practice, the biological definition of race has been difficult to apply to human populations. We decided to present these definitions in quotation to avoid unintentional misleading interpretations of them. That said, the importance of these is that they offer an idea of the complexity behind the discussion over human race. In this regard, several experts (e.g., Donovan et al., 2019; Jackson & Depew, 2017; Lieberman, 1968; Norton et al., 2019) reiterate that the debate over human race should be guided by research evidence rather than by biased and naïve viewpoints. Likewise, Jackson and Depew (2017) stress that unfortunately this debate has been influenced by scientific racism—a form of racism that defends the destructive and fallacious assumption that biological race exists.

Having clarified that race is a vague term that has multiple meanings and that there is no consensus on its definition, in the next paragraphs we discuss the impact of this concept in society. We start by pointing out that much of this impact is explained by a continuous tension between two opposing views of race, namely essentialism and constructionism (also known as constructivism) (Morning, 2011). The genetic essentialism of race can be defined as a sociocognitive bias which assumes that the genes inherent in people make same race individuals physically and/or behaviorally uniform and people of different races physically and/or behaviorally different (Andreychik and Gill, 2015, as cited in Donovan et al., 2019). In constructionism, race is considered a social construction that has no basis in biological science. According to Morning (2011), in the constructivist view, the purpose of a racial classification is to serve as an instrument of power to create and perpetuate economically, politically, and socially human divisions. Within this view, Omi and Winant (2015) claim that the determination of racial categories has fueled both state-based and experiential racial politics and vice versa. Tawa (2020) contends that essentialist beliefs can lead to racist attitudes. As he explains, science courses are legitimate and desirable scenarios in which students should be provided with opportunities to enrich their understandings about the concept of race from a social constructionist perspective. Consequently, Donovan et al. (2020) insist that more efforts and resources should be invested in promoting anti-essentialist understandings of race. Also, they stress that students should be educated to refute essentialist thinking and reduce racist attitudes.

Considering the opposite definitions of essentialism and constructionism, the question then arises: is race biologically meaningful? Mukhopadhyay et al. (2014) point out

that the concept of race has historically been linked to the belief that there are “natural” divisions of the human species. Additionally, they explain that this naturalization of human division has been also associated with the idea that there are clear-cut, discrete, easily distinguishable, and homogeneous subgroups or races that make easy to classify human beings into these racial groups. With this in mind, it is worth adding here that evolutionary biologists acknowledge that biological differences between human groups exist, that these differences are continuous along geographic distance, but that attempting to categorize people into biological races simply does not make sense (Graves, 2005, as cited in Graves, 2010). By the same token, Donovan (2014) notes that “the nature of race is still debated in science because identifying human population structure and interpreting its meaning is not straightforward” (p. 466).

As stated by Snyder (1962), the idea of race appears to have emerged during the seventeenth century colonial expansion of Western European powers. The validity, definition, and social consequences of the concept of race have recurrently been debated by biologists and physical anthropologists (Lieberman et al., 1992). In a recent *Nature Cancer* editorial (2020), it is asserted that a problem that deserves more attention is racism, which is defined in the current article “as the emotional conviction that race and behavior are linked in heredity and that some races are superior to others” (Lieberman, 1968, p. 133). Racist attitudes have been justified through the idea that inequality is a natural product of human biological difference. More to the point, attitudes that sustain racial inequality tend to be perpetuated through the use of the genetic essentialism of race.

There is evidence of racist attitudes in Australia (Dobinson & Mercieca, 2020), Colombia (Soler Castillo, 2019; Viecco Garzón et al., 2017), Brazil (Garcia Castro & Abramovay, 2006), and the United States (Clark and Hurd, 2020; Donovan et al., 2019; Tawa, 2020), among other countries. In the case of the USA, it is alarming to see the exacerbation of these attitudes in times of Covid-19. As Clark and Hurd (2020) observe, in this country people who are Black, Indigenous, or people of color would suffer the most from the Covid-19 crisis. Furthermore, in a recent *Nature* editorial (2020), attention is focused on the use of the Covid-19 pandemic as a pretext for racist attacks against people of Asian descent around the world.

Kostas Kampourakis (2018), the past editor of *Science & Education*, argues that the fact that the term “race” has influenced various kinds of discriminative acts in the past is a strong reason to seriously consider abandoning this term. Also, recent literature alerts us to the need to explore pragmatic and effective ways to deal with the most profound obstacles used in education to challenge the genetic essentialism of race. In science education, for instance, one obstacle is that the treatment of race in some science textbooks influences inaccurate beliefs about racial difference (Beltrán Castillo, 2017; Donovan, 2015a, 2016; Willinsky 2020). Donovan et al. (2020) stress that a second obstacle is that the exploration of how belief in genetic essentialism among students differs after pedagogical intervention about different domains of genetics knowledge has been paid too little attention. Moreover, they claim that little is known about whether such differences reflect (or not) a change in beliefs about race. A third obstacle is discussed by Tawa (2020) who observes that to date there is little evidence about the impact of social constructionist race education on various types of racial beliefs. A fourth obstacle has to do with the fact that some science news and biology curricula perpetuate racial bias by unintentionally spreading inaccurate beliefs about genetic essentialism (Donovan, 2014, 2015b, 2016, 2017; Morin-Chassé, 2020; Tawa, 2020). In relation to this, Donovan et al. (2019) present the following critical appraisal: “a biology curriculum that perpetuates racial bias by unintentionally increasing

inaccurate beliefs about racial difference is inhumane because it harms those who suffer from racial discrimination” (p. 530).

Beckwith et al. (2017) have proposed the incorporation of dialogues into classroom or informal science education settings to critically explore genetics, ancestry, and race, and thus challenge the genetic essentialism of race. To be precise, they, as members of the Genetics and Society Working Group (GSWG) (genesandsociety.org), created and proposed a set of three dialogues as a didactic resource to foster informed views of the concept of race, and combat fallacies about the connection of science and race. These dialogues have been presented at the City of Cambridge’s annual Science Festival in April 2015 and again in April 2016. These have been also used in an undergraduate non-majors genetics course at Emerson College and in an undergraduate biology capstone course at Northeastern University. We decided to use the second dialogue: *Two scientists, two perspectives* (Beckwith et al., 2017, p. 531) in our dialogue-based CTCS, for six reasons.

The first reason is that the dialogue revolves around the controversy of whether or not the concept of race in humans is biologically meaningful. Second, the argumentative interaction between two scientists is the heart of the dialogue. Third, having emerged in the GSWG—an interdisciplinary group composed of scientists, students, and professionals trained in a variety of disciplines, including genetics, sociology, ethics, and the law—*Two scientists, two perspectives* is an holistic dialogue rigorously documented and conceived for all members of society. Fourth, for more than 20 years, the authors of this dialogue have collectively discussed research on human genetic variation and the social impact of such information. Much of the critical spirit of this collective discussion is embodied in this dialogue (Beckwith et al., 2017). Fifth, *Two scientists, two perspectives* is a dialogue of 2150 words. This really makes its incorporation feasible in a university science course. Moreover, Archila (2017, 2018) has reported that readings between 2000 and 3000 words are not too short and not too long, and consequently, they are well received by undergraduates. The sixth and last reason is that this dialogue has never been used as a pragmatic and effective means to explicitly introduce undergraduates to the identification and the evaluation of arguments.

Two scientists, two perspectives (Beckwith et al., 2017) can briefly be described as follows:

[Jon and Tobi] are two scientists [who] explore the question of whether or not human race is biologically meaningful, tracing some of the history of alternative perspectives in their field. In the 18th and for much of the 19th century, scientists relied on physical traits as a means of identifying human racial groups. With the development of DNA technology, studies of the genetic basis of “race” became much more sophisticated, and differences between people from different parts of the world could now be roughly quantified. Jon argues that “race” has no biological basis. Tobi disagrees, arguing that there are distinct biological races that can be separated on the basis of physical characteristics and genetic information (p. 528).

4.1 The Dialogue-Based CTCS

Our dialogue-based CTCS is inspired in a drama-based CTCS created and proposed by Archila (2018). Nobel laureate Roald Hoffmann suggested to Archila that he could use the play *Should’ve* (Hoffmann, 2006) for educational purposes. Thus, Archila (2018) concluded that this play was an ideal educational resource for his drama-based CTCS in which 91 medical students in Colombia were asked to identify and evaluate arguments regarding

a dilemma between ethics, social responsibility, and scientific work presented in *Should've* (Hoffmann, 2006). In our case, we used *Two scientists, two perspectives* (Beckwith et al., 2017) in the dialogue-based CTCS as a springboard to articulate (1) argumentation, (2) critical thinking, (3) argument identification, (4) argument evaluation, (5) decision-making, and (6) controversial question (Fig. 1).

The scenario was designed as a single 75-min class session, and as recommended by Archila (2018), the teaching and learning conditions were to provide each student with opportunities to think critically to: (1) make a decision about a controversial question (“Is the concept of race applicable to humans?”), (2) identify, and (3) evaluate the arguments presented in the argumentative interaction between Jon and Tobi in the dialogue *Two scientists, two perspectives* (Beckwith et al., 2017). To be clear, identifying arguments is proposed as a transition phase between decision-making and argument evaluation (Fig. 1). This transition is based on the assumption that a good argument identification facilitates argument evaluation (Berland & Hammer, 2012, as cited in Archila, 2018).

We would like to stress that we consider that the major elements of our dialogue-based CTCS (Fig. 1) are pragmatic. This means that university science instructors can not only easily incorporate the classroom scenario into their educational practice, but also use it as a catalyst for small-group discussion and/or a whole-class debate, and thus take advantage of the highly controversial level of the question, “Is the concept of race applicable to humans?”. All this is possible because we created this scenario as an unfinished and open alternative for instructors interested in providing students with explicit opportunities to practice their argument identification and argument evaluation skills.

5 Method

5.1 Context and Participants

The dialogue-based CTCS was implemented in a university bilingual (Spanish–English) science course (Archila et al., 2018, 2021b) called Biology of Organisms. This course was chosen by *convenience sampling*—a sample that is selected because of its availability to the research team (Bryman, 2016). An important reason for this is that the second author is

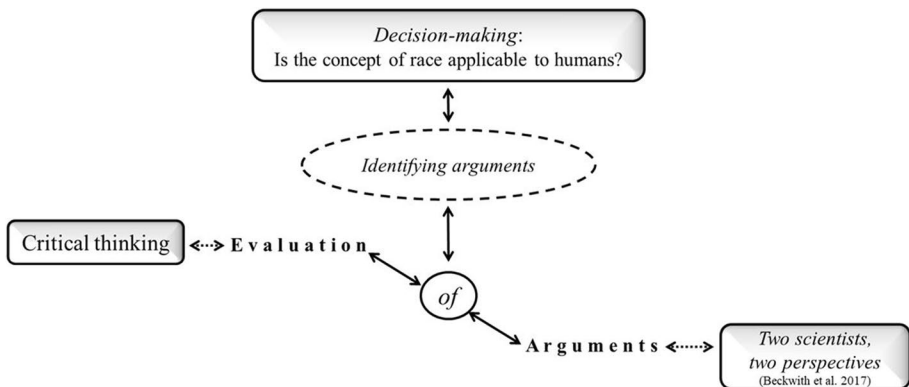


Fig. 1 Major elements of the dialogue-based CTCS (adapted from Archila, 2018, p. 57)

the course instructor. Biology of Organisms is a large (60–80 students per semester), introductory course that is offered every semester by the Department of Biological Sciences to participants in all undergraduate programs at a private university in Bogotá, Colombia. This university has a high academic ranking in Latin America. Its educational policy is to foster the integration of students from different majors and different age groups. Therefore, it is very common to see students from different socioeconomic status, academic achievement, majors (not only Biology and Microbiology), and ages taking this bilingual course.

Among the 131 eligible students, 117 (89.3%) participated in this study. Out of these 117 participants, 73 (62.3%) were females and 44 (37.6%) were males. The age distribution ranged from 16 to 24 years ($M_{\text{age}} = 18.6$, $SD_{\text{age}} = 1.74$). The authors informed the undergraduates that their answers would have no influence on their final course grade and that they could withdraw at any time. Participants and their parents were informed of the general research purpose. Permission to conduct this research was obtained from the University's Ethics Committee. All responses were kept confidential. The authors ensured that the inquiry was not harmful to any participant involved. All participants were treated in accordance with the ethical guidelines of the American Psychological Association with respect to consent, confidentiality, and anonymity. For this reason, the undergraduates were assigned codes to protect their privacy, for example 1U28 means Class 1, undergraduate number 28.

These 117 participants were grouped into two classes with the dialogue-based CTCS carried out in the following order:

Class 1: Undergraduates taking Biology of Organisms during the spring semester ($M_{\text{age}} = 18.5$), 63 students (35 females and 28 males).

Class 2: Undergraduates taking Biology of Organisms during the fall semester ($M_{\text{age}} = 18.7$), 54 students (38 females and 16 males).

It is important to clarify that in both classes, the CTCS was implemented at that stage of the course where the participants had the necessary subject matter knowledge required to understand the scientific content behind the controversy. To be clear, this matter knowledge was discussed in one unit of the syllabus of the Biology of Organisms, namely, "Evolution." This unit included key topics, such as allopatric and sympatric speciation, genetic variation, mechanisms of reproductive isolation, mechanisms of evolution, population genetics, and sexual selection. Also, it was more likely then that they would have the level of maturity required to understand the controversial issue, "Is the concept of race applicable to humans?"

5.2 Instructor Qualifications

As suggested by Archila (2018), to ensure university students were central in the scenario, the instructor who conducted the CTCS assumed the role of a facilitator. He is Full Professor and Head of the Department of Biological Sciences. He holds a Bachelor's degree in Biology, a Master's degree in Biological Sciences, and a Doctor rerum naturalium degree in neurobiology. He has been teaching the Biology of Organism course for more than 12 years. Recently, he has cooperated in various biology education innovation projects at university level. The instructor participated in the construction of the CTCS and was wholly committed to implementing the scenario, and perhaps more importantly, he introduced some changes in the Biology of Organism course for better implementation, such

as undergraduates' instruction in argumentation (how to identify and evaluate arguments) throughout the course. Prior to the implementation of this research project, the instructor had not met the student participants.

5.3 Research Design

It is important to clarify that to guarantee a transparent contrast of the results of our dialogue-based CTCS with those obtained by Archila (2018) in his drama-based CTCS, we decided to adopt the research design (data collection and data analysis) suggested by this academic. Accordingly, our study used a mixed methods approach where quantitative and qualitative analyses measures were used and considered complementary (Treagust et al., 2014, as cited in Archila, 2018). The data are the participants' responses to closed-ended (quantitative data) and open-ended questions (qualitative data) from a four-phase questionnaire (Appendix 1). It was felt that the quantitative findings could be useful to document achievements of our dialogue-based CTCS while the qualitative results would favor interpretation and description of participants' decisions, their argument identification, and evaluation.

5.3.1 Data Collection

Data were collected from the 117 participants' written responses. Participants were given a questionnaire (Appendix 1) by their instructor, who had been asked not to answer any questions except for clarification of terms. The questionnaire was composed of four phases. In the first and fourth phases, each undergraduate made their own decision about the controversy, "Is the concept of race applicable to humans?" The purpose of these phases was to have students make an initial (first phase) and a final (fourth phase) decision. In both phases, students' argumentation was sought through the question, "Why did you make that decision?"

In the second phase, students read the dialogue *Two scientists, two perspectives* (Beckwith et al., 2017). As mentioned previously, the argumentative interaction presented in the dialogue could be briefly summarized as two positions: (1) Jon argues that "race" has no biological basis, and (2) Tobi disagrees, arguing that there are distinct biological races that can be separated on the basis of physical characteristics and genetic information. In each class, during reading time, scenes were read aloud to the whole class by two students who assumed the roles of Jon and Tobi. Each student had a copy of the text to which s/he could refer during the CTCS. The purpose of this phase was to prompt participants to identify arguments. As the study by van Eemeren et al. (2015) recommends, participants were asked to "underline" (p. 736) (or highlight) in the text, Jon and Tobi's arguments (Appendix 1).

In the third phase, each participant was prompted to evaluate arguments by means of the following questions: (1) Are Jon's arguments solid? and (2) Are Tobi's arguments solid? Explanations for why or why not were required for each of the two questions (Appendix 1). Also, it should be pointed out that the questionnaire and the dialogue were distributed to the participants at the beginning of the CTCS. The whole questionnaire and the dialogue were in Spanish. Participants were given the option to decide the language (Spanish, English, or a hybrid version using code-switching) they wanted to use for writing their answers to each question due to the Spanish–English bilingual nature of the university science course (Archila et al., 2018, 2021b).

Finally, we considered that it was important to find out participants' opinions about the CTCS. Accordingly, at the end of the implementation, they were asked to answer a nine-item survey (5–10 min) (Appendix 2). Archila (2015a, 2015b, 2017) recommends generating a climate of confidence in the science classroom and asking students for key points that help the researchers continuously improve a pedagogical innovation. We created the survey based on questions previously formulated and documented by this researcher. For this reason, we assumed that these were valid to find out about participants' opinions of the CTCS. Participation in the survey was completely anonymous. The whole survey was in Spanish, and participants were given the option to decide which language (Spanish, English or a hybrid version using code-switching) they wanted to use to answer the open-ended questions.

5.3.2 Data Analysis

Analysis occurred at three levels: Analysis of participants' (1) initial and final decisions, (2) argument identification, and (3) argument evaluation. At the first level, participants' initial and final decisions ("Is the concept of race applicable to humans?") were classified: Yes/No. In their responses to the open-ended question, "Why did you make that decision?" the initial and the final decisions of each participant were compared and coded in order to determine whether the CTCS had enriched their argumentation (Table 1). The coding process was guided by the premise that argumentation requires the "justification of claims with reasons and/or evidence" (Erduran et al., 2020, p. 1). According to Erduran et al. (2020), an argument is valid when it is accompanied by reason (reason-based argument) and/or evidence (evidence-based argument). All data were coded and analyzed by the first author. A sample of the data (~50%) from the participants' initial and final decisions (Questions 2 and 8 in Appendix 1) was coded and analyzed by the second author. The coding was based on a single yes–no coding method and was conducted independently by the first and the second author. To assess the inter-coder reliability of this coding, the Cohen's kappa coefficient (Cohen, 1960) was carried out using the *Statistical Package for the Social Sciences* (SPSS®), and the coefficient was calculated as 0.74. According to Bryman (2016), values between 0.6 and 0.75 are considered a "good" (p. 276) inter-coder agreement. A consensus was reached after some discussion and further examination of students' responses.

At the second level, Jon and Tobi's arguments were identified by each participant and contrasted with a list of arguments identified by the second, third, and last authors, individually. After collating the arguments identified individually by these three authors, we consolidated a final list that included eighteen arguments (Jon, 5; Tobi, 13) (Appendix 3). At the third level, the evaluation of the argument responses ("Are Jon and Tobi's arguments solid?") (third phase of the questionnaire in Appendix 1) were classified as Yes/No. In addition, students' responses (Explain why or why not, Questions 5 and 6 in Appendix 1) were coded (Table 2). The purpose of this coding process was to determine whether the students effectively provided at least one reason why or why not they considered that the arguments put forth by Jon and Tobi were solid. As Davies and Barnett (2015), Ennis (2015), and Hitchcock (2017) remind us, the fact that individuals not only acknowledge the soundness of the arguments of others (even if they do not agree with them), but also communicate their reasons about their judgment, is a good indicator of fair-mindedness, which is an important feature (among others) of critical thinkers. All data were coded and analyzed by the first author. A sample of the data (~50%) was coded and analyzed by the second author. This process was based on a single yes–no coding method. The Cohen's

Table 1 Codes used in the coding data of the students' responses to the question, "Why did you make that decision?" Note: examples were taken from students' responses

Code	Description	Example
ANE = Argumentation not enriched	At least one new reason-based and/or evidence-based argument in the final decision is not found when compared with the initial decision	<p>At the beginning of the CTCS:</p> <p>"[Claim: I consider that the concept of race is not applicable to humans because] [Argument 1] race in humans is just an excuse to differentiate ourselves from each other based on our physical characteristics or geographic location. [Argument 2] In addition, a negative implication has been given to this term that even differentiates us by our capacities, socioeconomic status, and other characteristics" (2U6)</p> <p>At the end of the CTCS:</p> <p>"[Claim: I consider that the concept of race is not applicable to humans because] [Argument 1] the concept of race is a way of grouping and differentiating humans according to certain characteristics. [Argument 2] Moreover, race is a concept that can lead to discrimination" (2U6)</p>
AE = Argumentation enriched	At least one new reason-based and/or evidence-based argument in the final decision was found when compared with the initial decision	<p>At the beginning of the CTCS:</p> <p>"[Claim: I consider that the concept of race is applicable to humans because] [Argument 1] humans have distinctive physical characteristics that can be used to classify them" (2U18)</p> <p>At the end of the CTCS:</p> <p>"[Claim: I consider that the concept of race is not applicable to humans because] [Argument 1] there is no scientific basis to justify the use of certain physical and genetic characteristics to classify people. [Argument 2] It is also relevant to consider that a great number of the studies carried out to support this concept have been carried out by researchers belonging to one of these races, which also calls itself the superior race, there is an evident biased attitude on the part of these people. [Argument 3] Additionally, the use of physical characteristics, such as skin color to support the idea of human race is totally questionable. [Argument 4] Further, it is also common to find people appealing to a biased interpretation of results of the human genome to justify human divisions. [Argument 5] Moreover, the defenders of this concept not only lack reliable scientific evidence, but also manipulate the definition of race according to their own interests. [Argument 6] Finally, although the defenders of the concept of race argue that the classification of people is beneficial for the treatment of pathologies and the creation of public policies, the results show that this classification has had detrimental consequences on humanity. [Argument 7] Regarding the treatment of pathologies, there are environmental factors that have more scientific basis than that of supposed racial origin. [Argument 8] Additionally, the concept of race has simply caused racist policies which has been harmful to humanity" (2U18)</p>

Table 2 Codes used in the coding data of the students' responses to the question, "Are Jon and Tobì's arguments solid?" (Note: all eight examples were taken from students' responses)

Code	Description	Example	Comment
AENS = Argument evaluation not supported	The student did not provide at least one full reason to support her/his response to the question, "Are Jon and Tobì's arguments solid?" Explain why or why not	<p>Example 1: "[Jon's arguments are solid because] he uses inductive arguments and examples to refute Tobì's arguments." IU60</p> <p>Example 2: "[Jon's arguments are not solid because] he is not able to support his claims. His propositions are fallacious." IU27</p> <p>Example 3: "[Tobì's arguments are solid because] he gives convincing evidence and examples. He refutes Jon's arguments. His references are very logical. He is objective" (2U37)</p> <p>Example 4: "[Tobì's arguments are not solid because] these rely on studies influenced by subjectivity" (2U5)</p>	<p>This example serves to illustrate that the reason provided by IU60 to support her/his response is incomplete because s/he did not refer to at least one of the arguments and/or examples used by Jon</p> <p>Example 2 shows that IU17 did not clarify why s/he considered that Jon was not able to support his claims. Also, IU17 did not explain why s/he concluded that Jon's propositions were fallacious</p> <p>In this example, 2U37 provided an incomplete reason, since s/he did not give details of at least one piece of evidence and/or example used by Tobì</p> <p>Example 4 shows that 2U5 missed specifying at least one study mentioned by Tobì which 2U5 considered that was "influenced by subjectivity"</p>
AES = Argument evaluation supported	The student provided at least one full reason to support her/his response to the question, "Are Jon and Tobì's arguments solid?" Explain why or why not	<p>Example 5: "[Jon's arguments are solid because] he demonstrates that there is no robust biological basis for classifying people into races and shows the inconsistencies that scientists have to use to support this hypothesis. First, the way people are grouped according to their skin color, or their genes is arbitrary. Second, policies such as slavery, discrimination, eugenics, and imperialism were supported by the idea of the immutability of skills and behaviors predetermined by race. Furthermore, the scientists who created these classifications made racist claims confirming their lack of impartiality. In fact, the classification of people according to DNA sequencing of genomes has fanciful criteria that allow it to conveniently conform to the theory of the five races. This leads to no consistent definition of what race really is [...]" (1U50)</p> <p>Example 6: "[Jon's arguments are not solid because] he supports his argumentation only from a social view. For example, he mentions that social policies may be based on wrong assumptions of the concept of race and that this is destructive for society [...]" (2U30)</p>	<p>This example illustrates a full reason because IU50 not only provided at least one reason to support her/his response, but also gave specific details about this reason</p>
			<p>Example 6 shows that 2U30 produced at least one full reason since s/he included one concrete example</p>

Table 2 (continued)

Code	Description	Example	Comment
		<p>Example 7: “[Tobi’s arguments are solid because] he relies primarily on evidence from studies that have been verified by other scientists. Further, he highlights the possible benefits of classifying humans into five races (e.g., treatment of pathologies). He also constructs a strong argument related to the fact that scientists cannot be judged on the misuse of their theories” (2U22)</p>	<p>Here, 2U22 put forward specific information for at least one reason to support her/his response</p>
		<p>Example 8: “[Tobi’s arguments are not solid because] he asserts that classifying humans in different racial groups is a well-founded idea. However, he does not clarify under what criteria or what type of characteristics which can be used to group a number of people because a different number of groups can be determined. Besides, he is not able to explain the reasons for the genetic differences or similarities between people belonging to different geographical locations [...]” (1U38)</p>	<p>In this example, 1U38 presented in detail her/his reason why s/he considered that Tobi’s arguments were not solid</p>

kappa coefficient was calculated as 0.86 for Question 5 in Appendix 1 and 0.79 for Question 6 in Appendix 1. These values are considered a “very good” (Bryman, 2016, p. 276) inter-coder agreement. Any differences that emerged were discussed by the first and the second authors until a consensus was achieved. Finally, to enrich our analysis, participants’ responses to the survey (Appendix 2) were analyzed using frequency counts. Some answers to open-ended questions (Questions 4, 7, and 8 in Appendix 2) are commented on in the Results section.

6 Results

The results of the implementation of the dialogue-based CTCS are presented in the following three sections; the first section deals with decision-making (“Is the concept of race applicable to humans?”); the second section addresses the identification of the arguments of Jon and Tobi (Beckwith et al., 2017); and the third section focuses on argument evaluation (“Are Jon and Tobi’s arguments solid?”). The outcomes of the nine-item survey (Appendix 2) are presented throughout these three sections to provide a deeper understanding of the contribution of this research study.

6.1 Decision-Making

In this research, the dialogue *Two scientists, two perspectives* (Beckwith et al., 2017) was the backbone of a dialogue-based CTCS that revolved around the thought-provoking question, “Is the concept of race applicable to humans?” At the beginning as well as at the end of the scenario, each participant made a decision about this question (first and fourth phases in Appendix 1). Table 3 shows the participants’ responses to this controversial question.

The results displayed in Table 3 indicate that in both classes the participants are undoubtedly divided between “yes” and “no” answers. This corroborates the controversial nature of this question. Another key finding here is that in both classes the decisions made at the beginning (40 “yes” and 23 “no” in Class 1; 33 “yes” and 21 “no” in Class 2) were almost the same at the end (39 “yes” and 24 “no” in Class 1; 29 “yes” and 25 “no” in Class 2) of the dialogue-based CTCS. In our question, as in many other controversial questions, there was no straightforward yes/no answer: what was most important were the arguments constructed by the participants. Specifically, we expected that at the beginning (first phase) participants made a decision (“Is the concept of race applicable to humans?”) and communicated their arguments (“Why did you make that decision?”) based on their previous knowledge. And then, the identification (second phase) and the evaluation (third phase) of Jon and Tobi’s arguments helped them to enrich the arguments they communicated at the end (fourth phase) of the CTCS. After comparing the participants’ arguments constructed at the beginning and at the end of the CTCS, Table 4 shows the number of undergraduates that at the end stayed with the same arguments and the number of participants that enriched their arguments. A representative number of participants (52/63 in Class 1; 49/54 in Class 2) enriched their arguments at the end of the scenario. This is a promising outcome because as we have been emphasizing throughout this article one of the goals of higher education should be to promote argumentation rather than impose (indoctrinate) what students decide or believe.

Table 3 Decisions made at the beginning and at the end of the CTCS and by each class: “Is the concept of race applicable to humans?”

	At the beginning (<i>N</i> = 117)	%	At the end (<i>N</i> = 117)	%
Class 1 (<i>n</i> = 63)				
Yes	40	63	39	62
No	23	37	24	38
Class 2 (<i>n</i> = 54)				
Yes	33	61	29	54
No	21	39	25	46
Classes 1 and 2 (<i>N</i> = 117)				
Yes	73	62	68	58
No	44	38	49	42

6.2 Argument Identification

After making an initial decision (at the beginning of the dialogue-based CTCS), participants were asked to read the dialogue *Two scientists, two perspectives* (Beckwith et al., 2017) and identify (highlight) the arguments of Jon and Tobi (second phase in Appendix 1). Specifically, this dialogue recreates the argumentative interaction between Jon and Tobi. The former uses five arguments to support his claim that the concept of race is not applicable to humans while the latter communicates thirteen arguments to refute this claim (Appendix 3). Our results reveal that the number of Jon’s arguments identified by participants ranged from two to five ($M=4.2$, $SD=0.86$) in Class 1 and from one to five ($M=3.6$, $SD=1.24$) in Class 2, while the number of Tobi’s arguments identified by the students ranged from one to thirteen ($M=8.3$, $SD=2.69$) in Class 1 and from two to thirteen ($M=6.7$, $SD=2.79$) in Class 2. These outcomes indicate that in both classes, participants did not have much trouble in identifying at least one of the arguments set out by Jon and Tobi.

To better interpret these outcomes, the survey shows that a considerable number of participants (49/63 in Class 1; 42/54 in Class 2), apart from the Biology of Organisms course, had received instruction in argument identification (Question 1 in Appendix 2). In addition, a high number of undergraduates (55/63 in Class 1; 48/54 in Class 2) considered that the argument identification was useful to help them to make a decision (Question 7 in Appendix 2). Some of their reasons include the following: “this helped me to clarify the two positions [referring to Jon and Tobi] and their different arguments” and “this was useful to understand and analyze more easily and promptly the arguments of both people.” Arguably, these reasons illustrate the key role of argument identification in the decision-making process.

6.3 Argument Evaluation

After identifying Jon and Tobi’s arguments, undergraduate students were asked to evaluate arguments based on the question, “Are Jon and Tobi’s arguments solid?” (the third phase in Appendix 1). According to Erduran et al. (2020), an argument is solid (sound, valid) when it is accompanied by reason and/or evidence. It is important to bear in mind that the arguments presented by Jon and Tobi are all solid (Beckwith et al., 2017). Consider the following examples to better illustrate this. Example 1: Jon considers that the

Table 4 Number of students that enriched their arguments at the end of the CTCS

	At the end, stayed with the same arguments	At the end, enriched their arguments
Class 1 ($n=63$)	11	52
Class 2 ($n=54$)	5	49
$N=117$		

concept of race in humans is not biologically meaningful because “sequencing the DNA of human genomes showed that two people picked from any place in the world share about 99% of their DNA sequences” (Beckwith et al., 2017, p. 533). This argument is solid (valid) as it is accompanied by evidence (Erduran et al., 2020). Example 2: Tobi maintains that the concept of race in humans is biologically meaningful because “the 5 racial groups are Africans, Native Americans, Eurasians, East Asians, and Pacific Islanders. They differ in characteristics such as skin color, bone structure, physique etc., all traits with a genetic basis” (Beckwith et al., 2017, p. 532). This argument is solid as it is justified by reason (Erduran et al., 2020). Table 5 displays the results of students’ evaluation of arguments.

In both classes, more than half of the participating students considered that Jon’s (46/63 students in Class 1; 35/54 students in Class 2) and Tobi’s (39/63 students in Class 1; 33/54 students in Class 2) arguments were solid. In our CTCS, argument evaluation was created as a valuable opportunity for students to think critically about the concept of race in humans. Accordingly, they were required to support their answers (yes/no) with reasons (“explain why or why not”). Table 6 shows that a considerable proportion of the students gave at least one full reason to support their evaluation of Jon’s (51/63 students in Class 1; 44/54 students in Class 2) and Tobi’s (50/63 students in Class 1; 42/54 students in Class 2) arguments. This demonstrates that these participants were able to evaluate arguments as fairly as possible. It is important to note that this is a positive result as it means that students strived to conduct an unbiased evaluation of arguments, avoiding being dominated by the motivation to prioritize personal beliefs (Davies & Barnett, 2015; Ennis, 2015; Hitchcock, 2017).

The results of the survey indicate that more than half of the participants (38/63 in Class 1; 33/54 in Class 2), apart from the Biology of Organisms course, had received instruction in argument evaluation (Question 2 in Appendix 2). Moreover, a representative number of participants (62/63 in Class 1; 48/54 in Class 2) considered that the argument evaluation was useful to help them to make a decision (Question 8 in Appendix 2). Some comments include the following reasoning: “when I evaluated each argument, I analyzed and thought about how I would have had supported it if I was Jon or Tobi. This served me to generate my own arguments and ideas” and “it helped me to develop arguments to support my views. It was not only the quantity of arguments [Jon put forth five arguments and Tobi thirteen] but the quality of these.” These comments reinforce the importance of providing students with opportunities to evaluate arguments.

In our CTCS, the dialogue *Two scientists, two perspectives* (Beckwith et al., 2017) was used as a didactic tool to introduce undergraduates to the identification and the evaluation of arguments. The results of the survey show that nearly all the students (62/63 in Class 1; 51/54 in Class 2) mentioned they had sufficient time for reading this dialogue (Question 5 in Appendix 2). Some participants (49/63 in Class 1; 36/54 in Class 2) also commented that the reading aloud activity contributed to their understanding of this dialogue (Question

Table 5 Record of participants that' evaluation of arguments: "Are Jon and Tobi's arguments solid?"

	Jon (N=117)	%	Tobi (N=117)	%
Class 1 (n=63)				
Yes	46	73	39	62
No	17	27	24	38
Class 2 (n=54)				
Yes	35	65	33	61
No	19	35	21	39
Classes 1 and 2 (N=117)				
Yes	81	69	72	62
No	36	31	45	38

Table 6 Results of the participants' reasons as support to their evaluation of Jon and Tobi's arguments

	Jon				Tobi			
	AENS (N=117)	%	AES (N=117)	%	AENS (N=117)	%	AES (N=117)	%
Class 1 (n=63)	12	19	51	81	13	21	50	79
Class 2 (n=54)	10	19	44	81	12	22	42	78
Classes 1 and 2 (N=117)	22	19	95	81	25	21	92	79

AENS argument evaluation not supported, AES argument evaluation supported

6 in Appendix 2). Another interesting finding related to this dialogue is that a relevant number of respondents (55/63 in Class 1; 53/54 in Class 2) considered that it was easily understandable for them (Question 4 in Appendix 2). Some comments include the following: "its vocabulary did not contain strange words," "as this was a dialogue, the arguments were explicitly presented," and "it is easy to realize that the dialogue presents two opposite characters, their positions as well as their arguments." We assume these comments as feedback about the utility of the dialogue *Two scientists, two perspectives* (Beckwith et al., 2017) to provide opportunities for the students to think critically about the controversy over human race. We provided these opportunities in a context in which nearly all the participants (60/63 in Class 1; 51/54 in Class 2), apart from the Biology of Organisms course, had received instruction in critical thinking (Question 3 in Appendix 2). The point of concern here is that many undergraduates (51/63 in Class 1; 47/54 in Class 2) never (27/63 in Class 1; 22/54 in Class 2) or infrequently (24/63 in Class 1; 25/54 in Class 2) had had the opportunity to think critically about whether the concept of race is applicable to humans in other university courses (Question 9 in Appendix 2).

7 Discussion

To repeat, the preparation of an ideal "critical thinker" is a more complex (and ambitious) process than engaging students in argument identification and argument evaluation activities. It is therefore vital to clarify that even though we obtained promising results, we were not able to prepare ideal "critical thinkers," but this has never been the intention of our

study. Our intention was only aimed at reporting on a CTCS based on the controversy over human race as a pragmatic and effective means to introducing undergraduates to the identification and the evaluation of arguments, which are expressly promulgated by several scholars as two crucial elements, among many, characterizing critical thinking (e.g., Andrews, 2015; Archila, 2018; Epstein, 2017; Harker, 2015; Hitchcock, 2017; Jiménez-Aleixandre & Puig, 2012; Lau, 2011; Walton, 2019). The importance of this controversy is that it has been widely recognized by scholars as an authentic open-ended issue without a clear-cut solution due to its intrinsic relationship with multiple factors, such as ethical, moral, and political (Donovan, 2014, 2015a, 2015b, 2016, 2017; Donovan et al., 2019, 2020; Jackson & Depew, 2017; Morning, 2011; Mukhopadhyay et al., 2014; Tawa, 2020). Three research questions guided this exploratory investigation. In this section, the results are discussed in the light of these questions and previous studies. Emphasis is placed on the role of the dialogue *Two scientists, two perspectives* (Beckwith et al., 2017) in our dialogue-based CTCS.

First Research Question—“Does the CTCS provide participants with opportunities to make a decision about the controversial question of whether or not the concept of race in humans is biologically meaningful?”. The results of this study showed that the activity provided participants with opportunities to make an initial and a final decision (Table 3). In both classes, there was not a drastic change between the initial and the final decision. To be precise, the discrepancy among students about whether or not the concept of race in humans is biologically meaningful stayed until the end of the CTCS. It is perhaps most important to highlight the fact that our outcomes suggested that the scenario helped participants to enrich their argumentation independently of the decision they made (Table 4). This is key point since Andrews (2015) reminds us that one of the goals of higher education should be to foster argumentation rather than indoctrinate what students decide or believe. Therefore, our results reinforce Erduran et al.’s (2020) idea that helping students to become aware of the importance of justifying their claims with reasons and/or evidence is a step in the right direction, towards the cultivation of critical thinking skills.

Previous research has underscored the controversial nature of issues such as climate change (Rehg, 2011), evolution and creationism (Archila & Molina, 2020), GMO crops (Fahnestock, 2020), and nuclear energy (Jho et al., 2014). Our results expand these issues. In particular, we provide research evidence for the claim that the issue over human race is a powerful and effective source of controversy among university students, which can be used to introduce them to the identification and the evaluation of arguments. Moreover, our findings are in line with previous research which confirms that the use of socioscientific issues is beneficial for students’ engagement in argumentation practices such as the identification and the assessment of arguments (Evagorou & Osborne, 2013; Knight & McNeill, 2015).

Some academics maintain that more efforts and resources should be invested in creating effective and pragmatic activities focused on the cultivation of critical thinking skills (e.g., Andrews, 2015; Archila et al., 2019; Hyttinen et al., 2019; Osborne, 2010). Clearly, (1) effectiveness and (2) pragmatism are two relevant conditions to help instructors assume their role of facilitators of critical thinking and avoid the (mis)conception that conceptual content coverage of the syllabus should be the priority. The controversy over human race was a key aspect of our dialogue-based CTCS. The fact that making decisions about whether or not the concept of race in humans is biologically meaningful has caused discrepancy among undergraduate students in an introductory science course, which demonstrates that this CTCS is not only effective and pragmatic, but also realistic. This is a valuable finding as Morning (2011) mentions that “institutions of higher education are important

not just for what they impart to students, but for the influence they exercise on their faculty members' understandings of racial difference" (p. 46). In relation to this, the results of the survey (Appendix 2) showed that many undergraduates never or infrequently had the opportunity to think critically about whether the concept of race is applicable to humans in other university courses. Arguably, this offers an idea of how relevant our dialogue-based CTCS was.

Second Research Question—"Does the CTCS provide participants with opportunities to identify the arguments presented in a dialogue in which two scientists interact argumentatively to explore the controversial question?". The results indicate that in both classes students did not have much trouble in identifying at least one of the arguments communicated by Jon and Tobi. To be more precise, Jon put forth five arguments and Tobi thirteen (Appendix 3). Hence, the maximum possible arithmetic means for each were 5 and 13, respectively. In the case of Jon's arguments, we obtained mean of 4.2 in Class 1 and 3.6 in Class 2. And in relation to Tobi's arguments, a mean of 8.3 in Class 1 and 6.7 in Class 2 was found. There are at least two plausible reasons to explain these positive outcomes. First, the majority of participants, apart from the Biology of Organisms course, had previously received instruction in argument identification (Question 1 in Appendix 2). This means that, before this course, they had certain knowledge about how to identify arguments. Second, a representative number of students considered that the dialogue *Two scientists, two perspectives* (Beckwith et al., 2017) was easily understandable for them (Question 4 in Appendix 2). This demonstrates that this dialogue is an effective didactic resource to provide students with opportunities to identify arguments both for and against the human race controversy. In addition, it could be speculated that, as van Eemeren et al. (2015) argue, the fact that we have asked students to "underline" (p. 736) or highlight could have helped them to better identify arguments in the text. Furthermore, our results confirm those made by earlier researchers (Archila, 2018; van Eemeren et al., 2015) who dealt with the use of text as didactic resource for the advancement of argument identification.

Third Research Question—"Does the CTCS provide participants with opportunities to evaluate the arguments presented in a dialogue in which two scientists interact argumentatively to explore the controversial question?". The results suggest that in our CTCS, participants were provided with opportunities to evaluate Jon and Tobi's arguments (Table 5). Furthermore, nearly all the participants considered that the argument evaluation was useful in helping them make a decision about the controversial question, "Is the concept of race applicable to humans?" (Question 8 in Appendix 2). Argument assessment is closely related to evidence assessment, and both are complex activities (Archila, 2015a, 2015b). In the context of our CTCS, this complexity is portrayed in the fact that the assessment of the arguments presented by Jon and Tobi forced participants to analyze whether or not the articulation of Jon's (the concept of race is not applicable to humans) and Tobi's (the concept of race is applicable to humans) claims with the piece of evidence used by Jon and Tobi was solid within the argumentation process of these two people. This is a relevant outcome once it is recognized that students cultivate their critical thinking skills when they are involved in argument evaluation practices (Osborne, 2012). Moreover, this result expands on the scope of some notable work carried out previously that has examined the complexity of argument evaluation (Botting, 2016; Jørgensen, 2007; Selinger, 2014; Walton, 2015, 2016).

Donovan et al. (2020) and Tawa (2020) invite us to assume science courses as educational scenarios to combat racial bias and provide students with explicit opportunities to

reflect on the controversy over human race in an unbiased atmosphere. In a similar vein, Archila (2018) notes that fair-mindedness is one characteristic, among many others, of critical thinkers. Hence, a promising finding of our study is that more than half of the participants considered that Jon and Tobi's arguments were solid. These participants evaluated Jon and Tobi's arguments impartially—not based on pre-conceived prejudices. Key evidence of this is that many produced at least one complete reason to support their argument evaluation (Table 6). This means that they were respectful of Jon and Tobi's alternative viewpoints and saw both sides of the controversy over human race independently of their own points of view. Likewise, this result is in alignment with academics' call to pass from rhetoric to concrete actions to foster fair-mindedness, among other principles of critical thinking (Davies & Barnett, 2015; Ennis, 2015; Hitchcock, 2017).

In sum, the results (Tables 3, 4, 5, and 6) are experimental evidence that lead us to claim that the dialogue-based CTCS provided university students with useful opportunities to identify and evaluate arguments both for and against, and to make informed decisions about whether or not the concept of race in humans is biologically meaningful. In general, our outcomes confirm that the identification and the assessment of arguments are good supports to guide students in cultivating critical thinking skills useful to make decisions about controversial issues (Andrews, 2015; Epstein, 2017; Harker, 2015; Hitchcock, 2017; Jiménez-Aleixandre & Puig, 2012; Lau, 2011; Walton, 2019). In particular, the results reported in this article corroborate the importance and usefulness of including socioscientific issues as a way to give students explicit opportunities to reflect on the moral, ethical, and political factors that influence these issues and thus become aware of why there is no straightforward solution for them (Evagorou & Osborne, 2013; Knight & McNeill, 2015).

7.1 Limitations

Four notable limitations of the present study should be acknowledged. First, one hundred and seventeen participants is a relatively small sample size. Therefore, great caution needs to be taken regarding the generalizability of the results. Second, we adopted a “convenience sampling” (Bryman, 2016, p. 187) approach because of its accessibility to the research team. It would be interesting to implement our dialogue-based CTCS using another approach, such as “purposive sampling” (Bryman, 2016, p. 410). Clearly, this would require participating students with different assessments of biological knowledge, and/or race conceptions (essentialist, constructionist), and/or argumentation skills at the outset. This approach would offer several possibilities. For example, the study of the impact of the CTCS on students' race conceptions. Third, this study has been carried out just in one introductory Biology course in a Colombian university. Clearly, the implementation of the dialogue-based CTCS in other courses and other countries would be important to generate more robust evidence. The last serious limitation that should be considered is that half of the participating students, apart from the Biology of Organisms course, had previously received instruction in argument identification, critical thinking, and argument assessment. This is a contextual factor that may have influenced the outcomes obtained in this study as well as the fact that the great majority of participants never or infrequently had had the opportunity to think critically about the controversy over human race in other university courses.

8 Implications and Scope for Future Research

There are two implications that can be derived from the results of this study. First, if introducing undergraduates to the identification and the evaluation of arguments is as relevant to university science education as this article has claimed, instructors should be guided and supported in the creation, implementation, and evaluation of effective and pragmatic CTCSs. Second, a vital condition for CTCSs becoming more customary educational practices is that higher education institutions need to abandon the outmoded (mis)conception that the conceptual content coverage of the syllabus is the priority. This implies drastic but desirable changes, for example, to switch from instructor-centered learning to student-centered learning (Andrews, 2015; Wendland et al., 2015). We consider that the dialogue-based CTCS reported in this article is a concrete and realistic step in the right direction towards the construction of student-centered learning scenarios (Hoidn & Klemenčič, 2021).

As Archila (2018) notes, a CTCS is in essence an open and unfinished creation for instructors interested in providing students with opportunities to cultivate their critical thinking skills. In this sense, future avenues of investigation might enrich and adapt our dialogue-based CTCS to the circumstances of the practices of other instructors, other educational contexts, other dialogues, and other disciplines (e.g., biomedical engineering, chemistry), and other controversial issues. Also, future research could focus on the use of this CTCS as a springboard to engage students in other legitimate and desirable educational practices, such as instructor-student and student–student argumentative interaction in the form of small-group and whole class debates. This is a feasible possibility because the dialogue *Two scientists, two perspectives* (Beckwith et al., 2017) recreates an argumentative interaction between two scientists (Jon and Tobi). We recommend especially that future studies should be inspired by the premise that “thinking critically is a defense against a world of too much information and too many people trying to convince us” (Epstein, 2017, p. 1).

Appendix 1. Questionnaire

Phase One: Decision-making

1. In your opinion, is the concept of race applicable to humans?
2. Yes
3. No

Why did you make that decision?

Phase Two: Argument Identification

Use the text of the dialogue, *Two scientists, two perspectives* to complete the following activity.

3. Highlight in the text, Jon’s arguments.
4. Highlight in the text, Tobi’s arguments.

Phase Three: Argument Evaluation

In your opinion, are Jon's arguments solid? Explain why or why not.

In your opinion, are Tobi's arguments solid? Explain why or why not.

Phase Four: Decision-making

In your opinion, is the concept of race applicable to humans?

- a. Yes
- b. No

8. Why did you make that decision?

Appendix 2. Survey

1. Apart from the Biology of Organisms course, have you ever received instruction in argument identification?
2. Yes.
3. No.
4. Apart from the Biology of Organisms course, have you ever received instruction in argument evaluation?
5. Yes.
6. No.
7. Apart from the Biology of Organisms course, have you ever received instruction in critical thinking?
8. Yes.
9. No.
10. Was the dialogue, *Two scientists, two perspectives* easily understandable for you? Explain why or why not.
11. Did you have sufficient time to read the dialogue?
12. Yes.
13. No.
14. Did the reading aloud activity contribute to your understanding of the dialogue?
15. Yes.
16. No.
17. Was the argument identification useful for you to make a decision? Explain why or why not.
18. Was the argument evaluation useful for you to make a decision? Explain why or why not.
19. How often do you have the opportunity to think critically about whether the concept of race is applicable to humans in other university courses?
20. Very frequently.
21. Fairly frequently.
22. Infrequently.
23. Never.

Appendix 3. List of arguments identified by the second, third, and last authors in the dialogue *Two scientists, two perspectives* (Beckwith et al., 2017)

Jon's arguments

1. "There have been horrendous consequences of social policies that were based on assumptions that people are born with different skills and behaviors that cannot be changed. This kind of thinking helped support such destructive social policies as slavery, discrimination, eugenics, and imperialism" (p. 532).
2. "For instance, in the eighteenth century European scientist Johann Blumenbach studied race and concluded that white Caucasians were the most beautiful of races, and Linnaeus, a scientist active around the same time, defined Africans as the most capricious of races" (p. 532).
3. "Agassiz had no evidence for his ideas other than his own personal negative feelings about black people. Likewise, Morton's opinion that, quote, "blacks are the nearest approximation to the lowest animals" makes it highly unlikely that he was able to study human races in an unbiased fashion. Remember that Morton's work was used as support for slavery in the South" (p. 532).
4. "Sequencing the DNA of human genomes showed that two people picked from any place in the world share about 99% of their DNA sequences" (p. 533).
5. "Rather than looking at what environmental factors (for example, systematic discrimination, poverty, etc.) might cause a phenotype like high blood pressure, you try to see it all from a misleading genetic perspective" (p. 534).

Tobi's arguments

1. "The 5 racial groups are Africans, Native Americans, Eurasians, East Asians, and Pacific Islanders. They differ in characteristics such as skin color, bone structure, physique etc., all traits with a genetic basis" (p. 532).
2. "Races may differ in their likelihood of getting certain diseases" (p. 532).
3. "Some trace its beginnings to the eighteenth century when European scientists came up with ideas about the 5 classical races. Later, U.S. scientists contributed further support for the theories" (p. 532).
4. "Harvard Professor Louis Agassiz, in the nineteenth century, proposed that Africans and whites had evolved separately. And Dr. Samuel Morton in Philadelphia measured the skull capacity of different supposed racial types. His estimations of brain size supported the race theories of his predecessors" (p. 532).
5. "At any rate, with new tools and more rigor, 20th- and 21st-century scientists have continued to provide support for the conclusions about races we have been talking about" (p. 532).
6. "Most of these 1% differences in DNA are found in all human populations and geographic places. However, each population differs in what percentage of the people in that population have the variants" (p. 533).
7. "These frequencies were studied by genome scientists who developed a computer program to sort people into different groups using these similarities and differences in their DNA sequences. The computer program was able to take DNA sequence data

- and neatly group people into their geographical regions of origin based on those differences. And, the computer reported that it could sort the populations into 5 groups, corresponding to the geographical regions of the 5 classical races” (p. 533).
8. “The scientists were able to ask that the computer tell them what the groups would look like when the sequences are analyzed to divide people into 2, 3, 4, 5, or 6 groups. When they asked for 5 groups, the computer program reported those 5 geographical races that fit best. Other numbers yielded different groupings with more or fewer peoples classed together” (p. 533).
 9. “There have been subsequent studies based on more sequences that neatly split people into 18 groups by geographical regions” (p. 533).
 10. “In this study, 14 groups were within Africa and the 4 remaining groups within the rest of the world. So I would say that there are still the 5 racial types, but maybe the African race can be further divided into 14 groups within Africa alone” (p. 533).
 11. “The use lies in the fact that genetically defined races might have common features that are potentially useful to know” (p. 534).
 12. “Many African Americans with origins in West Africa have a strong tendency to have high blood pressure. Researchers found a mutation in a gene that is responsible for the condition” (p. 534).
 13. “Just African Americans seem to exhibit high blood pressure. Those tested in West Africa who had the same susceptibility gene did not have high blood pressure” (p. 534).

Acknowledgements The authors would like to thank all the undergraduate participants for their cooperation in this project. This study was funded by the Vice-Presidency of Research and Creation, Universidad de los Andes, Bogotá, Colombia.

Declarations

Conflict of interest The authors have no conflicts of interest to disclose.

References

- Andrews, R. (2015). Critical thinking and/or argumentation in higher education. In M. Davies & R. Barnett (Eds.), *The Palgrave handbook of critical thinking in higher education* (pp. 729–780). Palgrave Macmillan.
- Andreychik, M. R., & Gill, M. J. (2015). Do natural kind beliefs about social groups contribute to prejudice? Distinguishing bio-somatic essentialism from bio-behavioral essentialism, and both of these from entitativity. *Group Processes & Intergroup Relations*, *18*(4), 454–474.
- Archila, P. A. (2014a). *Comment enseigner et apprendre chimie par l'argumentation?* Éditions Universitaires Européennes.
- Archila, P. A. (2014b). Are science teachers prepared to promote argumentation? A case study with pre-service teachers in Bogotá city. *Asia-Pacific Forum on Science Learning and Teaching*, *15*(1), 1–21.
- Archila, P. A. (2015a). Using history and philosophy of science to promote students' argumentation. A teaching-learning sequence based on the discovery of oxygen. *Science & Education*, *24*(9), 1201–1226.
- Archila, P. A. (2015b). Evaluating evidence from a historical chemical controversy: A study in French high school. *Asia-Pacific Forum on Science Learning and Teaching*, *16*(2), 1–22.
- Archila, P. A. (2017). Using drama to promote argumentation in science education: The case of “Should've.” *Science & Education*, *26*(3–4), 345–375.
- Archila, P. A. (2018). Evaluating arguments from a play about ethics in science: A study with medical learners. *Argumentation*, *32*(1), 53–76.

- Archila, P. A., & Molina, J. (2020). Evolution and creationism: Views of students in a Colombian university—Findings from 7 years of data using a three-question survey. *Research in Science Education*, 50(4), 1619–1638.
- Archila, P. A., Molina, J., & Truscott de Mejía, A.-M. (2018). Using formative assessment to promote argumentation in a university bilingual science course. *International Journal of Science Education*, 40(13), 1669–1695.
- Archila, P. A., Molina, J., & Truscott de Mejía, A.-M. (2019). Promoting undergraduates' awareness of the importance of thinking critically about false or inaccurate scientific information presented in news articles. *Revista Eureka Sobre Enseñanza y Divulgación De Las Ciencias*, 16(13), 1–27.
- Archila, P. A., Molina, J., Danies, G., Truscott de Mejía, A.-M., & Restrepo, S. (2021a). Providing undergraduates with opportunities to explicitly reflect on how news articles promote the public (mis)understanding of science. *Science & Education*, 30(2), 267–291.
- Archila, P. A., Molina, J., & Truscott de Mejía, A.-M. (2021b). Fostering bilingual written scientific argumentation (BWSA) through collaborative learning (CL): Evidence from a university bilingual science course. *International Journal of Science Education*, 43(1), 1–29.
- Barnshaw, J. (2008). Race. In R. T. Schaefer (Ed.), *Encyclopedia of race, ethnicity, and society* (pp. 1091–1093). Sage.
- Beckwith, J., Bergman, K., Carson, M., et al. (2017). Using dialogues to explore genetics, ancestry, and race. *The American Biology Teacher*, 79(7), 525–537.
- Beltrán Castillo, M. J. (2017). Racismo científico y textos escolares de ciencias naturales (1979–2015). *Voces y Silencios. Revista Latinoamericana de Educación*, 8(1), 37–59.
- Berland, L. K., & Hammer, D. (2012). Students' framings and their participation in scientific argumentation. In M. S. Khine (Ed.), *Perspectives on scientific argumentation: Theory, practice and research* (pp. 73–93). Springer.
- Botting, D. (2016). The logical evaluation of arguments. *Argumentation*, 30(2), 167–180.
- Bryman, A. (2016). *Social research methods* (5th ed.). Oxford University Press.
- Clark, U. S., & Hurd, Y. L. (2020). Addressing racism and disparities in the biomedical sciences. *Nature Human Behaviour*, 4(8), 774–777.
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(1), 37–46.
- Davies, M., & Barnett, R. (2015). Introduction. In M. Davies & R. Barnett (Eds.), *The Palgrave handbook of critical thinking in higher education* (pp. 1–25). Palgrave Macmillan.
- Dobinson, T., & Mercieca, P. (2020). Seeing things as they are, not just as we are: Investigating linguistic racism on an Australian university campus. *International Journal of Bilingual Education and Bilingualism*. <https://doi.org/10.1080/13670050.2020.1724074>
- Donovan, B. M. (2014). Playing with fire? The impact of the hidden curriculum in school genetics on essentialist conceptions of race. *Journal of Research in Science Teaching*, 51(4), 462–496.
- Donovan, B. M. (2015a). Reclaiming race as a topic of the U.S. biology textbook curriculum. *Science Education*, 99(6), 1092–1117.
- Donovan, B. M. (2015b). Putting humanity back into the teaching of human biology. *Studies in History and Philosophy of Biological and Biomedical Sciences*, 52, 62–75.
- Donovan, B. M. (2016). Framing the genetics curriculum for social justice: An experimental exploration of how the biology curriculum influences beliefs about racial difference. *Science Education*, 100(3), 586–616.
- Donovan, B. M. (2017). Learned inequality: Racial labels in the biology curriculum can affect the development of racial prejudice. *Journal of Research in Science Teaching*, 54(3), 379–411.
- Donovan, B. M., Semmens, R., Keck, P., et al. (2019). Toward a more humane genetics education: Learning about the social and quantitative complexities of human genetic variation research could reduce racial bias in adolescent and adult populations. *Science Education*, 103(3), 529–560.
- Donovan, B. M., Weindling, M., & Lee, D. M. (2020). From basic to humane genomics literacy. How different types of genetics curricula could influence anti-essentialist understandings of race. *Science & Education*, 29(6), 1479–1511.
- Ennis, R. H. (2015). Critical thinking: A streamlined conception. In M. Davies & R. Barnett (Eds.), *The Palgrave handbook of critical thinking in higher education* (pp. 31–47). Palgrave Macmillan.
- Epstein, R. L. 2017. *Critical thinking*. Socorro: Advanced Reasoning Forum.
- Erduran, S., & Jiménez-Aleixandre, M. P. (2007). Argumentation in science education: An overview. In S. Erduran & M. P. Jiménez-Aleixandre (Eds.), *Argumentation in science education: Perspectives from classroom-based research* (pp. 3–27). Springer.
- Erduran, S., Guilfoyle, L., & Park, W. (2020). Science and religious education teachers' views of argumentation and its teaching. *Research in Science Education*. <https://doi.org/10.1007/s11165-018-9758-z>

- Evagorou, M. (2011). Discussing a socioscientific issue in a primary school classroom: The case of using a technology-supported environment in formal and nonformal settings. In T. Sadler (Ed.), *Socioscientific issues in the classroom* (pp. 133–160). Springer.
- Evagorou, M., Jiménez-Aleixandre, M., & Osborne, J. (2012). ‘Should we kill the Grey Squirrels?’ ‘A study exploring students’ justifications and decision-making. *International Journal of Science Education*, *34*(3), 401–428.
- Evagorou, M., & Osborne, J. (2013). Exploring young students’ collaborative argumentation within a socio-scientific issue. *Journal of Research in Science Teaching*, *50*(2), 209–237.
- Fahnestock, J. (2020). Rhetorical citizenship and the science of science communication. *Argumentation*, *34*(3), 371–387.
- García Castro, M., & Abramovay, M. (Eds.). (2006). *Relações raciais na escola: reprodução de desigualdades em nome da igualdade*. Brasília: Edições UNESCO.
- Graves, J. (2005). *The race myth: Why we pretend race exists in America*. Dutton Books.
- Graves, J. (2010). Biological V. social definitions of race: Implications for modern biomedical research. *The Review of Black Political Economy*, *37*(1), 43–60.
- Greco Morasso, S. (2009). The argumentum experience. In N. Muller Mirza & A.-N. Perret-Clermont (Eds.), *Argumentation and education: Theoretical foundations and practices* (pp. 215–235). New York: Springer.
- Harker, D. (2015). *Creating scientific controversies: Uncertainty and bias in science and society*. Cambridge University Press.
- Hitchcock, D. (2017). *On reasoning and argument*. Springer.
- Hoidn, S., & Klemenčič, M. (2021). Introduction and overview. In S. Hoidn & M. Klemenčič (Eds.), *The Routledge international handbook of student-centered learning and teaching in higher education* (pp. 1–13). Routledge.
- Hoffmann, R.. (2006). *Should’ve*. {Debió ser} (Miguel Delgado, Graciela Diaz de Delgado and Daniel Delgado Diaz, Trans.). Estado Mérida: Consejo de Publicaciones ULA.
- Hyytinen, H., Toom, A., & Shavelson, R. J. (2019). Enhancing scientific thinking through the development of critical thinking in higher education. In M. Murtonen & K. Balloo (Eds.), *Redefining scientific thinking for higher education* (pp. 59–78). Palgrave Macmillan.
- Jackson, J. P., & Depew, D. J. (2017). *Darwinism, democracy, and race. American anthropology and evolutionary biology in the twentieth century*. London: Routledge.
- Jho, H., Yoon, H.-G., & Kim, M. (2014). The relationship of science knowledge, attitude and decision making on socio-scientific issues: The case study of students’ debates on a nuclear power plant in Korea. *Science & Education*, *23*(5), 1131–1151.
- Jiménez-Aleixandre, M. P., & Puig, B. (2012). Argumentation, evidence evaluation and critical thinking. In B. J. Fraser, K. Tobin, & C. J. McRobbie (Eds.), *Second international handbook of science education* (pp. 1001–1015). Springer.
- Jørgensen, C. (2007). The relevance of intention in argument evaluation. *Argumentation*, *21*(2), 165–174.
- Kampourakis, K. (2018). Darwinism, democracy, and race. *Science & Education*, *27*(5–6), 589–590.
- Knight, A. M., & McNeill, K. L. (2015). Comparing students’ individual written and collaborative oral socioscientific arguments. *International Journal of Environmental & Science Education*, *10*(5), 623–647.
- Lau, J. Y. F. (2011). *An introduction to critical thinking and creativity: Think more, think better*. John Wiley & Sons.
- Lieberman, L. (1968). The debate over race: A study in the sociology of knowledge. *Phylon*, *29*(2), 127–141.
- Lieberman, L., Hampton, R. E., Littlefield, A., & Hallead, G. (1992). Race in biology and anthropology: A study of college texts and professors. *Journal of Research in Science Teaching*, *29*(3), 301–321.
- Marks, J. (2010). Ten facts about human variation. In M. P. Muehlenbein (Ed.), *Human evolutionary biology* (pp. 265–276). Cambridge University Press.
- McNeill, K. L. (2011). Elementary students’ views of explanation, argumentation, and evidence, and their abilities to construct arguments over the school year. *Journal of Research in Science Teaching*, *48*(7), 793–823.
- Morin-Chassé, A. (2020). Behavioral genetics, population genetics, and genetic essentialism. *Science & Education*, *29*(6), 1595–1619.
- Morning, A. (2011). *The nature of race: How scientists think and teach about human difference*. University of California Press.
- Mukhopadhyay, C. C., Henze, R., & Moses, Y. T. (2014). *How real is race?* (2nd ed.). AltaMira Press.
- National Research Council. (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. The National Academies Press.

- Nature (Editorial). (2020). End coronavirus stigma now. *Nature*, 580(7802), 165.
- N Cancer (Editorial). (2020). Speaking up against inequity and racism. *Nature Cancer*, 1(6), 563–564.
- States, N. G. S. S. L. (2013). *Next Generation Science Standards: For states, by states*. The National Academies Press.
- Nielsen, J. A. (2012). Science in discussions: An analysis of the use of science content in socioscientific discussions. *Science Education*, 96(3), 428–456.
- Norton, H. L., Quillen, E. E., Bigham, A. W., Pearson, L. N., & Dunsworth, H. (2019). Human races are not like dog breeds: Refuting a racist analogy. *Evolution: Education and Outreach*, 12(17), 1–20.
- Oliveira, A., Akerson, V., & Oldfield, M. (2012). Environmental argumentation as sociocultural activity. *Journal of Research in Science Teaching*, 49(7), 869–897.
- Omi, M. A., & Winant, H. (2015). *Racial formation in the United States* (3rd ed.). Routledge.
- Osborne, J. (2010). Arguing to learn in science: The role of collaborative, critical discourse. *Science*, 328(5977), 463–466.
- Osborne, J. (2012). The role of argument: Learning how to learn in school science. In B. J. Fraser, K. Tobin, & C. J. McRobbie (Eds.), *Second international handbook of science education* (pp. 933–949). Springer.
- Rehg, W. (2011). Evaluating complex collaborative expertise: The case of climate change. *Argumentation*, 25(3), 385–400.
- Relethford, J. H. (2013). *The human species: An introduction to biological anthropology* (9th ed.). McGraw-Hill.
- Sakschewski, M., Eggert, S., Schneider, S., & Bögeholz, S. (2014). Students' socioscientific reasoning and decision-making on energy-related issues—Development of a measurement instrument. *International Journal of Science Education*, 36(14), 2291–2313.
- Sampson, V., & Schleigh, S. (2013). *Scientific argumentation in biology: 30 classroom activities*. Arlington: National Science Teachers Association.
- Selinger, M. (2014). Towards formal representation and evaluation of arguments. *Argumentation*, 28(3), 379–393.
- Snyder, L. L. (1962). *The idea of racialism*. D. Van Nostrand Company.
- Soler Castillo, S. (2019). “¡Mira, un negro!”. *Elementos para pensar el racismo y la resistencia*. Bogotá: Universidad Distrital Francisco José de Caldas.
- Swartz, R. J., & Perkins, D. N. (1990). *Teaching thinking: Issues and approaches*. Pacific Grove: Midwest Publications.
- Tawa, J. (2020). Does social constructionist curricula both decrease essentialist and increase nominalist beliefs about race? *Science & Education*, 29(6), 1513–1540.
- Treagust, D. F., Won, M., & Duit, R. (2014). Paradigms in science education research. In N. G. Lederman & S. K. Abell (Eds.), *Handbook of research on science education* (Vol. II, pp. 3–17). Routledge.
- van Eemeren, F. H., Grootendorst, R., & Meuffels, B. (2015). The skill of identifying argumentation. In F. H. van Eemeren (Ed.), *Reasonableness and effectiveness in argumentative discourse* (pp. 733–741). Springer.
- Viecco Garzón, M. C., Curiel Gómez, R. Y., & Muñoz Estrada, E. S. (2017). La diversidad cultural como operador pedagógico para la transformación del discurso racista en un referente social de construcción étnica incluyente. *AGLALA*, 8(1), 65–83.
- Walton, D. (2015). Some artificial intelligence tools for argument evaluation: An introduction. *Argumentation*, 30(3), 317–340.
- Walton, D. (2016). *Argument evaluation and evidence*. Springer.
- Walton, D. (2019). Argumentation schemes & their application in argument mining. In J. A. Blair (Ed.), *Studies in critical thinking* (pp. 177–211). Windsor: Windsor Studies in Argumentation.
- Wendland, M. W., Robinson, C., & Williams, P. A. (2015). Thick critical thinking: Toward a new classroom pedagogy. In M. Davies & R. Barnett (Eds.), *The Palgrave handbook of critical thinking in higher education* (pp. 153–168). Palgrave Macmillan.

Authors and Affiliations

Pablo Antonio Archila¹  · **Jorge Molina**² · **Giovanna Danies**³ · **Anne-Marie Truscott de Mejía**⁴ · **Silvia Restrepo**¹

Jorge Molina
jmolina@uniandes.edu.co

Giovanna Danies
g-danies@uniandes.edu.co

Anne-Marie Truscott de Mejía
atruscot@uniandes.edu.co

Silvia Restrepo
srestrep@uniandes.edu.co

¹ Vice-Presidency of Research and Creation, Universidad de los Andes, Cra 1 N° 18A-12, 111711 Bogotá, Colombia

² Department of Biological Sciences, Universidad de los Andes, Cra 1 N° 18A-12, 111711 Bogotá, Colombia

³ Department of Design, Universidad de los Andes, Cra 1 N° 18A-12, 111711 Bogotá, Colombia

⁴ School of Education, Universidad de los Andes, Cra 1 N° 18A-12, 111711 Bogotá, Colombia