Evaluating evidence from a historical chemical controversy: A study in a French high school

Pablo Antonio ARCHILA

LDAR (EA 4344), Physics Department Université Paris Diderot-Paris 7 PRES Sorbonne-Paris-Cité, FRANCE

E-mail: pabloantonioarchila@yahoo.fr

Received 27 Oct., 2015
Revised 21 Dec., 2015

Contents

Abstract

This paper addresses the importance of evaluating evidence for enriching critical thinking in the chemistry classroom. The purpose of the study was to examine the usefulness of a historical chemical controversy in promoting students' assessment of evidence. The investigation was conducted in a high school in Melun, France. 63 participants (24 females and 39 males aged 16–17 years) evaluated evidence relating to the polemical question of who discovered oxygen, with Carl Wilhelm Scheele (1742–1786), Joseph Priestley (1733–1804), and Antoine Laurent de Lavoisier (1743–1794) being possible contenders. This evidence was provided by the play

"Oxygen", written by Carl Djerassi and Nobel laureate Roald Hoffmann (2001a), and was classified as either experimentation in science or scientific communication. The findings indicate that this historical chemical controversy helped to raise the students' awareness of the essential role of evidence evaluation in the advancement of chemistry. Furthermore, they reveal that the participants evaluated "easier" evidence relating to experimentation in science rather than evidence relating to scientific communication. The main conclusion drawn from this study is that historical chemical controversies can enhance learners' assessment of evidence.

Keywords: chemistry education, critical thinking, evidence evaluation, history of chemistry.

Introduction

Enriching students' critical thinking skills should be an explicit aim in chemical education. One reason for this is that "students entering college chemistry courses typically have received little instruction or encouragement to practice critical thinking skills" (Kogut, 1996, p. 220). A second reason is that "maintaining social cohesion is often the imperative that channels learners working in groups, rather than critical thinking" (Taber, 2015, p. 19). Therefore, promoting critical thinking should be a research goal in chemical education (Zhou et al., 2012; Zoller & Pushkin, 2007). In this regard, Jiménez-Aleixandre and Puig (2012) point out that "evidence evaluation is an essential component of critical thinking" (p. 1002).

Evaluating evidence is part of the reasoning process in chemistry (Barke, Harsch & Schmid, 2012) and should be considered to be a component of chemical literacy. Evidence serves three purposes in the progress of this science: (i) it helps the chemist to better understand a chemical phenomenon; (ii) it helps to support or refute chemical laws, theories, models, etc.; and (iii) it plays a crucial role in making informed decisions involving chemical content, such as may be encountered in socio-scientific issues and scientific controversies. In order to fulfil these three purposes, there must be proper evaluation of evidence. Indeed, chemists continuously evaluate evidence. For example, they frequently assess evidence from infrared (IR) spectroscopy and nuclear magnetic resonance (NMR) spectroscopy to examine the molecular composition of substances. Analogously, citizens need to be prepared to evaluate and use evidence to make informed decisions in everyday situations that involve science, for instance issues to do with air quality, genetic engineering, human reproduction, and so forth.

In this research, a historical chemical controversy was used as a source of evidence. "It should be realized that the place of history is not only to make a conceptual point but also to introduce the humanistic element into the process of learning science" (Klassen & Froese Klassen, 2014, p. 1523); consequently, the central argument of this paper is that evidence evaluation by students can be promoted when the opportunities they have in the chemistry classroom are explicit and deliberately planned by the teacher. By adopting a concrete and specific approach, this investigation sought to address the following questions:

1. Could a historical chemical controversy be useful in promoting students' assessment of evidence relating to experimentation in science and scientific communication?
2. What are the opportunities for and obstacles to the use of historical chemical controversies for enriching students' assessment of evidence?

Literature review

Evaluation of evidence: contributions to the enhancement of critical thinking

In this paper, critical thinking is understood as "a cognitive activity, associated with using the mind. Learning to think in critically analytical and evaluative ways means using mental processes such as attention, categorization, selection, and judgement" (Cottrell, 2005, p. 1). Barak, Ben-Chaim and Zoller (2007) suggest that there are several notable advantages of science education for promoting critical thinking. In fact, they consider that this ability is decisive for citizens in modern life. Similarly, Kuhn (2005) asserts that the modern world is characterized by technical and social complexities. This panorama has prompted a heightened interest in critical thinking in certain education programs in Australia (e.g. ACARA, 2012), Canada (e.g. CMEC, 1997), China (e.g. Bing & Thomas, 2006; Leung, 1991; Lewin, 1987), Colombia (e.g. MEN, 2006), England (e.g. NCE, 2014), France (e.g. MENESE, 2012), Japan (e.g. MEXT, 2000), Spain (e.g. MEC, 2007) and the United States (e.g. AAAS, 1993; NGSS, 2013; NRC, 2012; Yager & Brunkhorst, 2014).

Critical thinking is a fundamental skill for students of the 21st century (Choi, Ko & Lee, 2015). In science education, the need for promoting this kind of thinking is largely justified by studies that have confirmed most of its potentialities. Some of these are:

- The stimulation of students' curiosity, interest and motivation (Kogut, 1996)
- The promotion of effective logical thinking among students (Koray & Köksal, 2009)
- The improvement of students' comprehension of the nature of science (Montgomery, 2009)
- The betterment of learners' conceptual understanding (Kogut, 1996; Zoller & Pushkin, 2007)
- The enrichment of learners' argumentation and evidence evaluation (Chang, 2007; Chang & Rundgren, 2010; Jiménez-Aleixandre & Puig, 2012; Pallant & Lee, 2015)

As mentioned earlier, this paper focuses on the relevance of evaluating evidence for enriching critical thinking in the chemistry classroom. At this point, it is important to clarify that "thinking critically does not mean questioning all data, evidence and experts, but rather developing criteria for evaluating them" (Jiménez-Aleixandre & Puig, 2012, p. 1012, italics added). Indeed, the fact that students evaluate evidence does not necessarily mean that they use it to argue in chemistry class. The use of evidence is another goal that is not discussed in this paper.

On the one hand, the evaluation of evidence is a cognitive process requiring class activities and laboratory work that have been specifically designed to this end. Nonetheless, some chemistry teachers have difficulty in coming up with evidence evaluation scenarios in the classroom (Barak, Ben-Chaim & Zoller, 2007). The main reason for this is that some teacher training programs do not teach future teachers to promote thinking abilities (Archila, 2014ab; Xie & So, 2012; Zhou et al., 2012). The situation is all the more complicated when some school teachers are forced to spend more time in the classroom, leaving them with much less time for other activities that are also important in the teaching and learning process (e.g. learning assessment, exchanging viewpoints with other teachers and parents, lesson preparation, reading students' production). This last has been strongly emphasized by Sahlberg (2010):

> Although […] teachers' work consists primarily of classroom teaching, many of their duties lay *outside of class*. […] in Finland […] teachers devote *less* time to teaching than do teachers in many other nations. For example, a typical middle school teacher in Finland teaches just less than 600 hours annually, corresponding to about four 45-minute lessons a day. In the United States, by

> contrast, a teacher at the same level devotes 1,080 hours to teaching over 180 school days […]. This, however, does not imply that teachers in Finland work less than they do elsewhere. An important—and still voluntary—part of […] teachers' work is devoted to the improvement of classroom practice, the school as a whole, and work with the community. Because […] teachers take on significant responsibility for curriculum and assessment, as well as experimentation with and improvement of teaching methods, some of the most important aspects of their work are conducted *outside of classrooms* (p. 7, italics added).

Teachers' actions should be regarded as crucial to the promotion of evidence evaluation (Jirout & Klahr, 2012; Koray & Köksal, 2009; Morris et al., 2012; Yenice, 2012). Nevertheless, one must bear in mind that the identity and the attributes shown by teachers in their practice are a complex construct, essential for encouraging professional development (Aristizábal, 2014; Enyedy, Goldberg & Welsh, 2006; Nogueira, 2014). These considerations would be highly relevant to initiatives in curriculum reform (Hernández, 2014).

On the other hand, Gott and Duggan (2003), Jiménez-Aleixandre and Puig (2012), Judge, Jones and McCreery (2009), Pallant and Lee (2015), and Yun and Kim (2015) emphasize that the evaluation of evidence can help individuals elaborate and communicate their own conceptions, opinions and postures. According to Gott and Duggan (2003), citizens are confronted by scientific evidence every single day; this might explain why it is imperative to offer learners opportunities to evaluate evidence in the chemistry classroom. Evidence evaluation is certainly useful for making decisions in issues such as climate change (Pallant & Lee, 2015), human cloning (Jiménez-Aleixandre & Puig, 2012) and methods of birth control (Gott & Duggan, 2003). In addition, Montgomery (2009) reports that evidence evaluation may assist in the development of a more informed understanding of the nature of science.

In this study, two types of evidence are discussed: (i) experimentation in science and (ii) scientific communication. The evidence that is produced during experiments helps chemists to better understand chemical phenomena (Lehman & Bensaude-Vincent, 2007). Additionally, scientists need to communicate their investigations. The existence of papers or books can be understood as evidence of scientific communication (Nielsen, 2013). Indeed, "there is an international push to improve the effectiveness with which scientists communicate" (Mercer-Mapstone & Kuchel, 2015, p. 1614).

The use of historical controversies as an educational tool in science

As mentioned in the introduction, this study concerns the use of a historical chemical controversy to promote students' assessment of evidence. This section presents an overview of past and current research involving this type of controversy. It is important to keep in mind that:

> If we wish to use the history of science to influence [enrich] students' understanding of science, we must include significant amounts of historical material and treat that material in ways which illuminate particular characteristics of science (Russell, 1981, p. 56).

This statement forms the basis of several studies that have explored the possible contributions of the history and philosophy of science (HPS) to the enhancement of learners' understanding of the nature of science (e.g. Abd-El-Khalick, 2013; Allchin, Andersen & Nielsen, 2014; Matthews, 1994). Some authors have called for further research to examine the conditions and impact of a combined approach between HPS and thinking abilities for facilitating argumentation (Adúriz-Bravo, 2014) and critical thinking (Montgomery, 2009). This call confirms that "HPS is being shown to be also relevant to problems in the learning of the sciences" (Matthews, 1994, p. 208).

The literature on HPS and chemistry teaching has continued to expand in recent years (e.g. Garritz, 2013; Greca & Freire, 2014; Niaz & Rodríguez, 2000). Some studies show efforts to place HPS on the curriculum (Monk & Osborne, 1997; Welch, 1979) and to explore the advantages for chemical education (Izquierdo, 2013). There is increasing research interest in the use of historical controversies in science education (Archila, 2015; de Hosson, 2011; Montgomery, 2009; Niaz, 2000, 2009). There is also a prevalent consensus on the need for more investigations that would provide empirical data for examining and expanding the scope of this type of controversy.

The historical content often included in science textbooks (Russell, 1981) and chemistry textbooks (Niaz & Coştu, 2013) needs to be dramatically improved. A historical and philosophical view of chemistry shows the existence of controversies as a habitual part of the daily work of a chemist (Garritz, 2013; Greca & Freire, 2014). Tsaparlis and Finlayson (2014) recognize that the use of historical chemical controversies is imperative in chemical education. Indeed, controversies in science are useful for helping students to understand how scientists actually work (Silverman, 1992). Hence, historical chemical controversies could offer opportunities for students to enrich their learning of chemistry and their learning about chemistry. Some examples of historical chemical controversies that have been used in education are listed below:

- Controversy between Jöns Jacob Berzelius's (1779–1848) and Auguste Laurent's (1807–1853) postulates over the substitution phenomenon (Archila, 2014c)
- Edward Turner's (1796–1837) postulate and the atomic weight controversy (Campbell, 1981)
- Controversy between the European school lead by Svante August Arrhenius (1859–1927) and the British school lead by Henry Edward Armstrong (1848–1937) over the dissociation phenomenon (de Berg, 2014)
- Controversy surrounding quantum mechanics and quantum chemistry (Garritz, 2013)
- Controversy between Robert Andrews Millikan (1868–1953) and Felix Ehrenhaft (1879–1952) surrounding "The oil drop experiment" (Niaz, 2000; Niaz & Rodríguez, 2005)

It is important to clarify that promoting evidence evaluation has not been the objective behind the inclusion of the aforementioned examples in textbooks. This is where the originality and authenticity of the present study come in. Our paper addresses the controversy between the chemists Carl Wilhelm Scheele (1742–1786), Joseph Priestley (1733–1804) and Antoine Laurent de Lavoisier (1743–1794) surrounding the question, "Who discovered oxygen?". This controversy is recreated in the play "Oxygen" written by Carl Djerassi and Nobel laureate Roald Hoffmann (2001a). "The ethical issues around priority and discovery at the heart of this play are as timely today as they were in 1777" (Djerassi & Hoffmann 2001b, p. 5). Hence, in this investigation drama is used as a learning strategy in science education (Braund, 2015; Klassen & Froese Klassen, 2014; Ødegaard, 2003; Pongsophon, Yutakom & Boujaoude, 2010). In addition, "drama is meant to be a very powerful teaching strategy for enhancing meaningful learning in science" (de Hosson & Kaminski, 2007, p. 622).

The literature review presented in this section is what inspired the combined HPS–critical thinking approach proposed in this study (Figure 1). The premise of this investigation is that a historical chemical controversy can provide students with evidence for evaluation. As mentioned earlier, evidence evaluation promotes critical thinking (Jiménez-Aleixandre & Puig, 2012) and a more informed understanding of the nature of science (Montgomery, 2009).

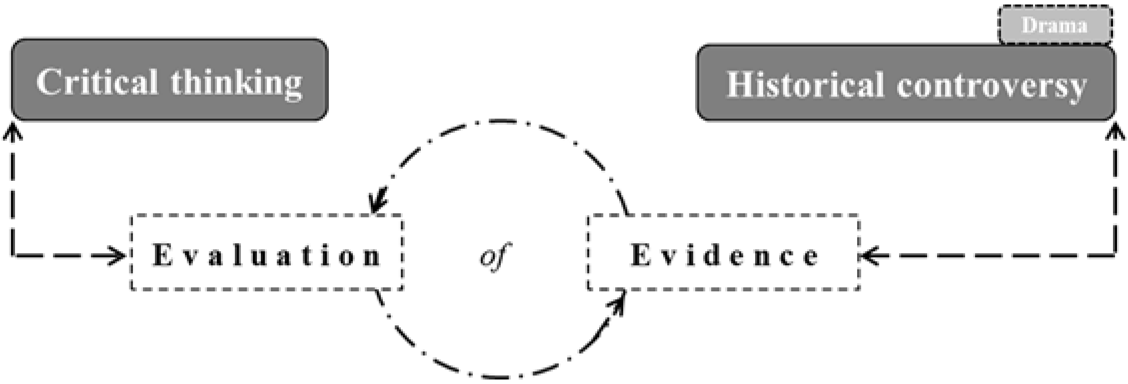

Figure 1. Combined approach between critical thinking and HPS.

Methodology

Research Design

This study used quantitative research design in a decision-making structure (Archila, 2015; Jho, Yoon & Kim, 2014; Maloney, 2007). This method was particularly useful to determine the percentage of participants considered each evidence provided by Scheele, Priestley and Lavoisier to be adequate in the controversy, “Who discovered oxygen?”. Another advantage of this method was the possibility of having more data about students’ assessment of evidence relating to experimentation in science and scientific communication.

Qualification of the instructor

The chemistry teacher is a high teacher with 32 years’ experience and with a strong background in chemistry and physics. She participates in various working groups at the academic level. When the data of this research were taken, she knew the 63 students for 4 months. It is important to keep in mind that the chemistry teacher in this study assumed the role of a facilitator; her sole function was to encourage the students.

Data collection

This study was conducted in a high school in Melun, France. The participants represented a wide range of achievement levels—low, medium and high, based on the chemistry grades of the previous semester. Data were collected from the written responses of 63 participants (24 females and 39 males aged 16–17 years) from households of middle-class socio-economic status. This data collection was conducted in two phases. In the first phase, the participants read Scene 8 of the French version of the play "Oxygen" (Djerassi & Hoffmann, 2003), which has an extension of 1650 words. The intervention of the chemists' wives was deleted to better focus the learners' attention on the experimentation in science and scientific communication recreated in this drama. Scene 8 can be briefly described as follows:

> The Chemistry Committee of the Royal Swedish Academy of Sciences decides to focus on the discovery of Oxygen, since that event launched the modern chemical revolution. But who should be so honored? Lavoisier is a natural choice, for if there ever was a marker for the beginning of modern chemistry, it was Lavoisier's understanding of the true nature of combustion, rusting, and animal respiration, and the central role of oxygen in each of these processes, formulated in the period 1770-1780. But what about Scheele? What about Priestley? Didn't they first discover oxygen?
>
> Indeed, on an evening in October 1774, Antoine Lavoisier, the architect of the chemical revolution, learned that the Unitarian English minister, Joseph Priestley, had made a new gas. Within a week, a letter came to Lavoisier from the Swedish apothecary, Carl Wilhelm Scheele, instructing the French scientist how one might synthesize this key element in Lavoisier's developing theory, the lifegiver oxygen. Scheele's work was carried out years before, but remained unpublished until 1777.
>
> Scheele and Priestley fit their discovery into an entirely wrong logical framework—the phlogiston theory—that Lavoisier is about to demolish. How does Lavoisier deal with the Priestley and Scheele discoveries? Does he give the discoverers their due credit? And what is discovery after all? Does it matter if you do not fully understand what you have found? Or if you do not let the world know?
>
> In a fictional encounter, the [Scene 8] brings the three protagonists [Carl Wilhelm Scheele, Joseph Priestley and Antoine Laurent de Lavoisier] and their wives [Sara Margaretha Pohl, Mary Priestley and Anne-Marie Pierrette Paulze Lavoisier] to 1777 Stockholm at the invitation of King Gustav III [...]. The question to be resolved: "Who discovered oxygen?" […]. In the Judgment of Stockholm, a scene featuring chemical demonstrations, the three discoverers of oxygen recreate their critical experiments (Djerassi & Hoffmann 2001b, pp. 4–5).

Scene 8 was chosen because in it, each of the three chemists faces continuous pressure to defend his title as the true discoverer of this gas. In this scene the evaluation of evidence offered by each chemist is crucial.

In the second phase, the participants answered the following questions:

1. Are Scheele's arguments adequate? Explain why or why not.
2. Are Priestley's arguments adequate? Explain why or why not.
3. Are Lavoisier's arguments adequate? Explain why or why not.

The chemistry teacher supervising the session did not participate in any way while the students were answering the three aforementioned questions. She took care not to influence their decisions by demonstrating and maintaining her neutrality. Thus, the students did not know *a priori* if the evidence presented by the three chemists was correct or not. Requiring the participants to support their answers with reasons ("explain why or why not") forced them to evaluate the evidence at hand.

Data analysis

The participants' responses constitute the data of this study. These responses were assumed to result from their evaluation of the evidence. All the data were analyzed by the author. It should be pointed out that all the evidence recreated in Scene 8 is correct (Bensaude-Vincent & Van Tiggelen, 2003). The author classified this evidence as two types: (i) experimentation in science and (ii) scientific communication. The aim of this classification was to provide criteria for identifying and classifying the participants' responses. The evidence corresponding to each of these types is presented below (Djerassi & Hoffmann, 2003):

Experimentation in science

Scheele's evidence

A. "Dissolve silver in acid of nitre and precipitate it with alkali of tartar. Wash the precipitate, dry it and reduce it by means of the burning lens in your apparatus […] A mixture of two airs will be emitted and pure silver left behind [talking to Lavoisier]" (p. 93).
B. "I did that experiment in 1771, in a pharmacy in Uppsala with equipment much more modest than now put at our disposal by your Majesty [King Gustav III]" (p. 95).

C. "I obtained the air over the next three years in many different ways, including red mercurius calcinatus, as you did [talking to Priestley]" (p. 95).

Priestley's evidence

D. "I made that air by heating mercurius calcinatus in 1774" (p. 92).
E. "In August of 1774, I exposed mercurius calcinatus, the red crust that forms as mercury is heated in air, in my laboratory to the light of my burning lens. As the red solid is heated, an air will be emitted, while dark mercury globules will condense on the walls of the vessel. You will collect the air by bubbling it though water. As soon as the gas appears […] catch it under water" (p. 96).

Lavoisier's evidence

F. "I have brought from Paris a suit of rubber I have devised. It catches all the effluents of the body to show us that the equation balances" (p. 99).
G. "I had begun my experiments with mercurius calcinatus …" (p. 102).

Scientific communication

Scheele's evidence

H. "In my book, about to appear" (p. 92).
I. "The experiment I brought to your attention some three years ago in my letter [talking to Lavoisier]" (p. 93).
J. "I told Professor Bergman [about the discovery]. I thought he would tell others" (p. 95).

Priestley's evidence

K. "I communicated that discovery in the same year [1774]" (p. 92).

Lavoisier's evidence

L. "[talking to Scheele] I know of no letter" (p. 93).

Findings

The findings are presented in two sections. The first section deals with the evaluation of evidence relating to experimentation in science, while the second concerns the evaluation of evidence relating to scientific communication.

Evaluation of evidence relating to experimentation in science

As mentioned earlier, all the evidence provided by Scheele, Priestley and Lavoisier in Scene 8 is correct and crucial (Bensaude-Vincent & Van Tiggelen, 2003). And yet, Table 1 shows that only 52% (33/63), 41% (26/63) and 48% (30/63) of the participants considered Evidence A, B and C, provided by Scheele, to be adequate in the controversy, "Who discovered oxygen?". Scheele's evidence has to do with the fact that he knew how to prepare the gas (Evidence A), produced the gas before Priestley and Lavoisier (Evidence B), and obtained the air over the next three years in many different ways (Evidence C).

Table 1. Evaluation of evidence relating to experimentation in science (N-C= non-classified answers)

	Experimentation in science
Scheele	
A	52% (33)
B	41% (26)
C	48% (30)
N-C	5% (3)
Priestley	
D	52% (33)
E	63% (40)
N-C	10% (6)
Lavoisier	
F	63% (40)
G	13% (8)
N-C	29% (18)

This paper does not deal with the "use of evidence" but "evidence evaluation". Nonetheless, the use of evidence depends on evidence evaluation (Jiménez-Aleixandre & Puig, 2012; Pallant & Lee, 2015). In other words, students should evaluate evidence carefully before using it. In this regard, Djerassi and Hoffmann (2001b) consider that there is no right answer to the question, "Who discovered oxygen?", pointing out that each chemist (Scheele, Priestley and Lavoisier) offers significant evidence. Despite Scheele's evidence being of significant value, the participants did not assess it, contrary to our expectations. Evidence B is significant in many ways in the controversy and could, for example, serve to prove that Scheele discovered oxygen. 41% (26/63) of the participants evaluated Evidence B in an appropriate manner (Table 1). This implies that the rest of the students (59%; 37/63) did not consider Scheele's producing of the gas before Priestley and Lavoisier (Evidence B) to be important evidence.

When it comes to Priestley, Scene 8 of the play "Oxygen" (Djerassi & Hoffmann, 2003) recreates two strong pieces of evidence in the controversy, "Who discovered oxygen?": he prepared the gas in 1774, three years after Scheele (Evidence D), and conducted a suitable chemical experiment to produce that gas (Evidence E). The participants' responses indicate that 52% (33/63) of them assessed Evidence D as being relevant (Table 1). Similarly, 63% (40/63) of the students recognized the significance of Evidence E.

Lavoisier's evidence was also evaluated by the students. According to Table 1, more than half of the participants (63%; 40/63) were aware of the importance of the "suit of rubber" (Evidence F), which Lavoisier invented in the eighteenth century to enable him to perform his experiment with greater care and exactitude. By contrast, only 13% (8/63) of the students considered Evidence G to be substantial in this controversy. This evidence confirms that Lavoisier knew—as did Scheele and Priestley—the method for preparing the gas through the decomposition of "mercurius calcinatus" (mercuric oxide, HgO). The results also indicate that 29% (18/63) of the learners did not assess Lavoisier's evidence (Table 1).

If all the evidence produced by Scheele, Priestley and Lavoisier in Scene 8 is correct and crucial (Bensaude-Vincent & Van Tiggelen, 2003), why is it that none of the evidence was evaluated by *all* (100%; 63/63) of the participants? The answer to this question is complex because, as the situation reveals, the participants struggled to properly assess evidence from the controversy. This observation expands on the findings of Xiao and Sandoval (2015) and Zoller and Pushkin (2007) regarding students' difficulty in evaluating evidence from socio-scientific issues and problem-based laboratory practice, respectively.

Evaluation of evidence relating to scientific communication

In this research, the books and papers that scientists publish are assumed to be a crucial part of scientific communication (Nielsen, 2013). In the controversy "Who discovered oxygen?" Scheele, Priestley and Lavoisier all provide decisive evidence of this type (Djerassi & Hoffmann, 2003).

Scheele provides three relevant facts. First, he wrote a book (which remained unpublished until 1777) to communicate his discovery (Evidence H). Second, at the end of 1774, he sent a letter to Lavoisier detailing his experiments (Evidence I). Third, he told Professor Bergman about his discovery, thinking that he would tell others (Evidence J, which is the least reliable). Table 2 shows that Evidence I (49%; 31/63) was more widely evaluated than either Evidence H (37%; 23/63) or Evidence J (19%; 12/63).

Table 2. Evaluation of evidence relating to scientific communication (N-C= non-classified answers)

	Experimentation in science
Scheele	
H	37% (23)
I	49% (31)
J	19% (12)
N-C	5% (3)
Priestley	
K	30% (19)
N-C	70% (44)
Lavoisier	
L	6% (4)
N-C	94% (59)

As for Priestley, he conducted his experiment on August 1, 1774. This chemist mentions something that should be understood as evidence relating to scientific communication: in October, he communicated his observations to Lavoisier (Evidence K). The results indicate that 70% (44/63) of the students had difficulty recognizing the relevance of Priestley's communication. A similar result was obtained with the situation in which Lavoisier was unaware of any letter from Scheele (Evidence L): very few of the students (6%; 4/63) considered this evidence to be decisive (Table 2). Yet, this evidence is undeniably important in the critical examination of "how much Lavoisier depended upon Priestley [and Scheele] for his understanding of oxygen" (Marshall & Marshall, 2005, p. 31).

A more detailed analysis of the results from Table 2 leads to the observation that the most widely evaluated evidence was Scheele's sending of a letter to Lavoisier in which he details his experiments (Evidence I, 49%; 31/63). By contrast, the fact that Lavoisier was unaware of the existence of a letter from Scheele (Evidence L) obtained the lowest evaluation. Evidence I and L were extremely decisive in that they led to the critical view that either Scheele (Evidence I) or Lavoisier (Evidence L) discovered oxygen. Indeed, both Evidence I and L are correct (Djerassi & Hoffmann, 2001b). However, the results indicate that the participants were not aware of their relevance.

The findings on the evaluation of evidence relating to scientific communication (Table 2) confirm the ongoing challenges of promoting the informed understanding of scientific communication as part of the nature of science. Nielsen (2013) claims that "it entails enabling science learners and teachers to observe scientific communication as an explicit and reflexive goal of science education itself" (p. 2080). More recently, Mercer-Mapstone and Kuchel (2015) declared the need for explicitly enhancing science communication skills. Consequently, the evaluation of evidence relating to scientific communication has the potential to not only help promote critical thinking but also to address the challenges reported by Nielsen (2013) and Mercer-Mapstone and Kuchel (2015).

Discussion

Are historical chemical controversies useful for promoting students' assessment of evidence relating to experimentation in science and scientific communication? This is the first question addressed in this study. To answer it, our investigation has used the eighteenth-century chemical controversy between Scheele, Priestley and Lavoisier recreated in Scene 8 of the play "Oxygen" (Djerassi & Hoffmann, 2003). According to the results (Tables 1 and 2), this historical controversy offers relevant information (evidence) that can be utilized in the chemistry classroom to promote learners' assessment of evidence.

Moreover, the results indicate that students are more inclined to evaluate evidence relating to experimentation in science (Table 1) than evidence relating to scientific communication (Table 2). Both types of evidence are decisive in the advancement of chemistry (Lehman & Bensaude-Vincent, 2007; Nielsen, 2013). Based on this finding, it is plausible to suggest that the chemistry classroom should become a place for enhancing students' assessment of evidence. More precisely, class activities should be planned to offer greater opportunities for learners to evaluate evidence *explicitly* in decision-making. This is an imperative condition for enriching evidence evaluation by students. Thus, the results of this research contribute toward the promotion of critical thinking by students through *supported argumentation* in the chemistry classroom itself. This is consistent with McDonald and McRobbie's (2012) reflections. According to these authors, it is important to:

> Also introduce the notion of *supported argumentation* instruction to describe an instructional approach to argumentation that does not explicitly guide learners in understanding the skills of argument, but instead provides prompts and suggestions for constructing arguments or *evaluating evidence* (p. 972, italics added).

What are the opportunities for and obstacles to the use of historical chemical controversies for enriching students' assessment of evidence? This is the second question addressed in this study. One obstacle to the promotion of critical thinking in the chemistry classroom is that pupils often make assertions without evidence (Kogut, 1996). This is a good reason to include evidence evaluation in the teaching of chemistry. Thus, if students evaluate evidence adequately, they would have more chances to use it. Often, as pointed out by Kogut (1996), Silverman (1992) and Yun and Kim (2015), learners consider that scientific questions necessarily trigger unambiguous, unique and correct answers. In this sense, the historical chemical controversy between Scheele, Priestley and Lavoisier used in this study shows that the assessment of evidence promotes a more informed understanding of the nature of science.

In the controversy, "Who discovered oxygen?", all the evidence provided by Scheele, Priestley and Lavoisier is correct and valid (Bensaude-Vincent & Van Tiggelen, 2003). In this study, the expectation was that the participants would consider most of the evidence provided by each chemist to be "decisive". However, the results suggest that students are not accustomed to evaluating evidence from historical chemical controversies. Scene 8 of the play "Oxygen" (Djerassi & Hoffmann, 2003) recreates valuable evidence that could serve to answer the question, "Who discovered oxygen?". For instance, Scheele prepared the gas in 1771 (before Priestley and Lavoisier) (Evidence B, 41%; 26/63), Priestley communicated his discovery in the same year that he performed his experiment

(Evidence K, 30%; 19/63), while careful weighings and experiments were integral in enabling Lavoisier to understand that the gas was oxygen (Evidence F, 63%; 40/63). Yet, none of this "decisive evidence" was widely evaluated by the participants (Tables 1 and 2).

Limitations of the Study

Evaluating evidence (Gott & Duggan, 2003; Jiménez-Aleixandre & Puig, 2012; Judge, Jones & McCreery, 2009; Pallant & Lee, 2015; Yun & Kim, 2015) and historical controversies (Archila, 2015; de Hosson, 2011; Montgomery, 2009; Niaz, 2000, 2009) are two research interest that have worked separately. In other words, this is the first study that conjugates them, not only to promote critical thinking, but also a more informed understanding of the nature of science. That said, various limitations must be addressed. It should be taken into consideration that in this study 63 participants (24 females and 39 males aged 16–17 years) evaluated evidence relating to the polemical question of who discovered oxygen. Thus, (1) the small number of participants, (2) the historical controversy chosen, and (3) the range of age are three strong limitations of the research reported in this paper.

Conclusion

This study is based in an unpublished approach combining evidence evaluation and a historical chemical controversy. That is why results are original. Evidence evaluation enhances critical thinking (Jiménez-Aleixandre & Puig, 2012; Kuhn, 2005). The overall conclusion that can be drawn from the results of this research is that historical controversies can be utilized to promote critical thinking in science education (Montgomery, 2009). The controversy "Who discovered oxygen?" appears to be particularly promising in terms of encouraging students to evaluate evidence relating to experimentation in science and scientific communication. There is, of course, much work to be done in this area, and further research must consider other historical controversies, other types of evidence, and other parts of the world.

Acknowledgements. The author gratefully acknowledges the contribution of Roald Hoffmann who provided the "Study Guide for Oxygen" frequently cited in this paper. The author also wishes to thank the participants of this study and their teacher, Catherine Royer.

References

Abd-El-Khalick, F. (2013). Teaching with and about nature of science, and science teacher knowledge domains. *Science & Education, 22*(9), 2087–2107.

Adúriz-Bravo, A. (2014). Revisiting school scientific argumentation from the perspective of the history and philosophy of science. In M. R. Matthews. (Ed.).*International handbook of research in history, philosophy and science teaching* (pp. 1443–1472). Dordrecht: Springer.

American Association for the Advancement of Science (AAAS). (1993). *Benchmarks for science literacy*. New York: Oxford University Press.

Allchin, D., Andersen, H. M., & Nielsen, K. (2014). Complementary approaches to teaching nature of science: Integrating student inquiry, historical cases, and contemporary cases in classroom practice. *Science Education, 98*(3), 461–486.

Archila, P. A. (2014a). 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. (2014b). Argumentation in chemistry teacher education: Past, present and future opportunities. *Revista Científica Vozes dos Vales, 6*,1–12.

Archila, P. A. (2014c). La argumentación en la formación de profesores de química: Relaciones con la comprensión de la historia de la química. *Revista Científica*. *18*, 50-66.

Archila, P. A. (2015). 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.

Aristizábal, C. A. (2014). Configuración de la identidad profesional docente como producto cultural. In A. Molina. (Ed.). *Enseñanza de las ciencias y cultura: Múltiples aproximaciones* (pp. 165–121). Bogotá: Ediciones Universidad Distrital Francisco José de Caldas.

Australian Curriculum, Assessment and Reporting Authority (ACARA). (2012). *Australian curriculum: Science F-10 version 3.0.* Sydney: Australian Curriculum, Assessment and Reporting Authority.

Barak, M., Ben-Chaim, D., & Zoller, U. (2007). Purposely teaching for the promotion of higher–order thinking skills: A case of critical thinking. *Research in Science Education, 37*(4), 353–369.

Barke, H.-D., Harsch, G., & Schmid, S. (2012). *Essentials of chemical education.* Heidelberg: Springer.

Bensaude-Vincent, B., & Van Tiggelen, B. (2003). Foreword. In C. Djerassi., & R., Hoffmann. *Oxygène : Pièce en 1 acte.* {Oxygen} (A, Kornprobst & J-C, Kornprobst, J-C., Trans. and Adap.). Toulouse : Presses Universitaires du Mirail.

Bing, W., & Thomas, G. P. (2006). An examination of the change of the junior secondary school chemistry curriculum in the PR China: In the view of scientific literacy. *Research in Science Education*, *36*(4), 403–416.

Braund, M. (2015). Drama and learning science: An empty space? *British Educational Research Journal, 41*(1), 102–121.

Campbell, W. A. (1981). Edward Turner (1796–1837) and the atomic weight controversy. *Analytical proceedings*, *18*(9), 381–383.

Chang, S.-N. (2007). Teaching argumentation through the visual models in a resource-based learning environment. *Asia-Pacific Forum on Science Learning and Teaching, 8*(1) 1–15.

Chang, S.-N., & Rundgren, C.-J. (2010). SEE-SEP: From a separate to a holistic view of socioscientific issues. *Asia-Pacific Forum on Science Learning and Teaching, 11*(1), 1–24.

Choi, Y., Ko, Y., & Lee, H. (2015, April). *Enhancing Korean middle school students' 21st century skills through collective intelligence based SSI instruction.* Paper presented at National Association for Research in Science Teaching (NARST) annual meeting in Chicago, US.

Cottrell, S. (2005). *Critical thinking skills. Developing effective analysis and argument*. New York: Palgrave Macmillan.

Council of Ministers of Education, Canada (CMEC). (1997). *Common framework of science learning outcomes K to 12*. Toronto: Council of Ministers of Education, Canada.

de Berg, K. C. (2014). The significance of the origin of physical chemistry for physical chemistry education: The case of electrolyte solution chemistry. *Chemistry Education Research and Practice, 15*(3), 266–275.

de Hosson, C. (2011). Una controversia histórica al servicio de una situación de aprendizaje: Una reconstrucción didáctica basada en diálogo sobre los dos máximos sistemas del mundo de Galileo. *Enseñanza de las Ciencias, 29*(1), 115–126.

de Hosson, C., & Kaminski, W. (2007). Historical controversy as an educational tool: Evaluating elements of a teaching–learning sequence conducted with the text "Dialogue on the ways that vision operates". *International Journal of Science Education, 29*(5), 617–642.

Djerassi, C., & Hoffmann, R. (2001a). *Oxygen*. Weinheim: Wiley–VCH.

Djerassi, C., & Hoffmann, R. (2001b). *Study guide for oxygen*. Weinheim: Wiley–VCH.

Djerassi, C., & Hoffmann, R. (2003). *Oxygène. Pièce en 1 acte*. {Oxygen} (A, Kornprobst & J-C, Kornprobst, J-C., Trans. and Adap.). Toulouse : Presses Universitaires du Mirail.

Enyedy, N., Goldberg, J., & Welsh, K. M. (2006). Complex dilemmas of identity and practice. *Science Education, 90*(1), 68–93.

Garritz, A. (2013). Teaching the philosophical interpretations of quantum mechanics and quantum chemistry through controversies. *Science & Education, 22*(7), 1787–1807.

Gott, R., & Duggan, S. (2003). Understanding and using scientific evidence: How to critically evaluate data. London: Sage.

Greca, I. M., & Freire, Jr. O. (2014). Teaching introductory quantum physics and chemistry: Caveats from the history of science and science teaching to the training of modern chemists. *Chemistry Education Research and Practice, 15*(3), 286–296.

Hernández, R. (2014). Contexto cultural y currículum en la enseñanza de las ciencias. In A. Molina. (Ed.). *Enseñanza de las ciencias y cultura: Múltiples aproximaciones* (pp. 145–163). Bogotá: Ediciones Universidad Distrital Francisco José de Caldas.

Izquierdo, M. (2013). School chemistry: An historical and philosophical approach. *Science & Education, 22*(7), 1633–1653.

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. G. Tobin., & C. J. McRobbie. (Eds.). *Second international handbook of science education* (pp. 1001–1015). Dordrecht: Springer.

Jirout, J., & Klahr, D. (2012). Children's scientific curiosity: In search of an operational definition of an elusive concept. *Developmental Review, 32*(2), 125–160.

Judge, B., Jones, P., & McCreery, E. (2009). *Critical thinking skills for education students.* Padstow: Learning Matters.

Klassen, S., & Froese Klassen, C. (2014). Science teaching with historically based stories: Theoretical and practical perspectives. In M. R. Matthews. (Ed.). *International handbook of research in history, philosophy and science teaching* (pp. 1503–1529). Dordrecht: Springer.

Kogut, L. S. (1996). Critical thinking in general chemistry. *Journal of Chemical Education, 73*(3), 218–221.

Koray, Ö., & Köksal, M. S. (2009). The effect of creative and critical thinking based laboratory applications on creative and logical thinking abilities of prospective teachers. *Asia-Pacific Forum on Science Learning and Teaching, 10*(1), 1–13.

Kuhn, D. (2005). *Education for thinking* . Cambridge, MA: Harvard University Press.

Lehman, C., & Bensaude-Vincent, B. (2007). Public demonstrations of chemistry in eighteenth century France. *Science & Education, 16*(6), 573–583.

Leung, J. Y. M. (1991). Curriculum development in the P. R. China. In C. J. Marsh., & P. Morris. (Eds.). *Curriculum development in East Asia* (pp. 61–81). London: Falmer.

Lewin, K. M. (1987). Science education in P. R. China: Transformation and change in the 1980's. *Comparative Education Review, 31*(3), 419– 441.

Maloney, J. (2007). Children's roles and use of evidence in science: An analysis of decision-making in small groups. *British Educational Research Journal, 33*(3), 371–401.

Marshall, J. L., & Marshall, V. R. (2005). Rediscovery of the elements: Joseph Priestley. *The Hexagon, 94*(2), 28-33.

Matthews, M. R. (1994). *Science teaching: The role of history and philosophy of science.* Routledge, New York.

Mercer-Mapstone, L. D., & Kuchel, L. J. (2015). Teaching scientists to communicate: Evidence-based assessment for undergraduate science education. *International Journal of Science Education, 37*(10), 1613–1638.

Ministère de l'Éducation nationale, de l'Enseignement supérieur et de la Recherche, France (MENESE). (2012). *School education in France*. Paris: Éduscol.

Ministerio de Educación Nacional, Colombia (MEN). (2006). *Estándares básicos de competencias en lenguaje, matemáticas, ciencias y ciudadanas.* Bogotá: Ministerio de Educación Nacional, Colombia.

Ministerio de Educación y Ciencia, Spain (MEC). (2007). *Real Decreto 1631/2006 Ensenanzas mínimas educación secundaria obligatoria.* Madrid: Ministerio de Educación y Ciencia, Spain.

Ministry of Education, Culture, Sports, Science and Technology, Japan (MEXT). (2000). *Education in Japan 2000: A graphic presentation.* Tokyo: Gyosei Corporation.

Montgomery, K. (2009). Using a historical controversy to teach critical thinking, the meaning of "theory", and the status of scientific knowledge. *Journal of Geoscience Education, 57*(3), 214–221.

Monk, M., & Osborne, J. (1997). Placing the history and philosophy of science on the curriculum: A model for the development of pedagogy. *Science Education, 81*(4), 405–424.

Morris, B. J., Croker, S., Masnick, A. M., & Zimmerman, C. (2012). The emergence of scientific reasoning. In H. Kloos., B. J. Morris., & J. L. Amaral. (Eds.). *Current topics in children's learning and cognition* (pp. 61–82). Rijeka: InTech.

National Curriculum in England (NCE). (2014). *Science programmes of study.* London: Department for Education.

National Research Council (NRC). (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas.* Washington, DC: National Academies Press.

Next Generation Science Standards (NGSS). (2013). *Next generation science standards: For states by states.* Washington, DC: National Academies Press.

Niaz, M. (2000). The oil drop experiment: A rational reconstruction of the Millikan–Ehrenhaft controversy and its implications for chemistry textbooks. *Journal of Research in Science Teaching, 37*(5), 480–508.

Niaz, M. (2009). Progressive transitions in chemistry teachers' understanding of nature of science based on historical controversies. *Science & Education, 18*(1), 43–65.

Niaz, M., & Coştu, B. (2013). Analysis of Turkish general chemistry textbooks based on a history and philosophy of science perspective. En M. S. Khine. (Ed.). *Critical analysis of science textbooks* (pp. 199-218). Dordrecht: Springer.

Niaz, M., & Rodríguez, M. A. (2000). Teaching chemistry as rhetoric of conclusions or heuristic principles – a history and philosophy of science perspective. *Chemistry Education Research and Practice, 1*(3), 315–322.

Niaz, M., & Rodríguez, M. A. (2005). The oil drop experiment: Do physical chemistry textbooks refer to its controversial nature? *Science & Education, 14*(1), 43–57.

Nielsen, K. H. (2013). Scientific communication and the nature of science. *Science & Education, 22*(9), 2067–2086.

Nogueira, L. (2014). O ensino de qualidade e as implicações na formação de professores. *Revista Científica Vozes dos Vales, 6*,1–14.

Ødegaard, M. (2003). Dramatic science. A critical review of drama in science education. *Studies in Science Education, 39*(1), 75–101.

Pallant, A., & Lee, H.-S. (2015). Constructing scientific arguments using evidence from dynamic computational climate models. *Journal of Science Education and Technology, 24*(2–3), 378–395.

Pongsophon, P., Yutakom., & Boujaoude, S. B. (2010). Promotion of scientific literacy on global warming by process drama. *Asia-Pacific Forum on Science Learning and Teaching, 11*(1), 1–38.

Russell, T. L. (1981). What history of science, how much, and why? *Science Education, 65*(1), 51–64.

Sahlberg, P. (2010). The secret to Finland's success: Educating teachers. *Stanford Center for Opportunity Policy in Education ~ Research Brief*, September, 1-8.

Silverman, M. P. (1992). Raising questions: Philosophical significance of controversy in science. *Science & Education, 1*(2), 163–179.

Taber, K. S. (2015). Meeting educational objectives in the affective and cognitive domains: Personal and social constructivist perspectives on enjoyment, motivation and learning chemistry. In M. Kahveci., & M. Orgill. (Eds.). *Affective dimensions in chemistry education* (pp. 3–27). Heidelberg: Springer.

Tsaparlis, G., & Finlayson, O. E. (2014). Physical chemistry education: Its many facets and aspects. *Chemistry Education Research and Practice, 15*(3), 257-265.

Welch, W. W. (1979). Twenty years of science curriculum development: A look back. *Review of Research in Education, 7*(1), 282–306.

Xiao, S., & Sandoval, W. A. (2015, Abril). *Functional roles of inscriptional evidence in children's written arguments about socioscientific issues.* Paper presented at National Association for Research in Science Teaching (NARST) annual meeting in Chicago, US.

Xie, Q., & So, W. (2012). Understanding and practice of argumentation: A pilot study with Mainland Chinese pre–service teachers in secondary science classrooms. *Asia–Pacific Forum on Science Learning and Teaching, 13*(2), 1–20.

Yager, R. E., & Brunkhorst, H. (Eds.). (2014). *Exemplary STEM programs: Designs for success*. Arlington, Virginia: NSTA Press.

Yenice, N. (2012). A review on learning styles and critically thinking disposition of pre-service science teachers in terms of miscellaneous variables. *Asia-Pacific Forum on Science Learning and Teaching, 13*(2), 1–31.

Yun, S. M., & Kim, H.-B. (2015). Changes in students' participation and small group norms in scientific argumentation. *Research in Science Education, 45*(3), 465–484.

Zhou, Q., Yan, C., Zhao, S., Liu, L., & Xing, L. (2012). A preliminary investigation into critical thinking of in-service and pre-service middle school chemistry teachers in Shaanxi province of China. *Asia-Pacific Forum on Science Learning and Teaching, 13*(2), 1–13.

Zoller, U., & Pushkin, D. (2007). Matching higher-order cognitive skills (HOCS) promotion goals with problem–based laboratory practice in a freshman organic chemistry course. *Chemistry Education Research and Practice, 8*(2), 153–171.