

The credibility of information on the Internet: criteria stated and criteria used by Mexican students^I

Marina Kriscautzky^{II}
Emilia Ferreiro^{III}

Abstract

The frequency of Internet use for purposes of study among the young and children highlights the need to use medium-specific criteria of credibility. The issue is of concern in education and has been investigated mainly by means of multiple choice surveys. The originality of this work is the design and implementation of a tool that allows approaching the inconsistencies between declarative data and data close to the decisions made when in action. TICómetro® is an online survey applied to 628 Mexican students, aged 14 to 18 years. We analyzed four of the twenty-six questions in the survey. One of them brings a multiple choice list of criteria for assessing the credibility of information; the other three questions require the selection of credible websites, based on images taken from the Internet (screen shots). Such procedure is closer to the actual search on the screen and it allows comparing the criteria stated to the criteria actually used. Statistical analysis of the data shows that the stated criteria do not match those used when selecting credible websites, in both age groups. However, there were significant differences between the two age groups with respect to the criteria. Another contribution of this study is the proposal of a new grouping of declarative data using the classical distinction between text and paratext, from literary theory, with the necessary adjustments to deal with informational texts in the digital space.

Keywords

Credibility of information – Digital information – Internet – Young students.

I- This article reports some of the results of the first author's PhD dissertation, which was developed in *Departamento de Investigaciones Educativas* (Department of Educational Research) of CINVESTAV, under the supervision of Dr. Emilia Ferreiro. We thank the authorities of Dirección General de Cómputo y de Tecnologías de Información y Comunicación (General Directorate of Computer and Information and Communication Technology) of UNAM and the entire staff of Coordinación h@bitat puma for making the implementation of TICómetro® possible.

II- Universidad Nacional Autónoma de México, Ciudad de México, México.
Contact: mkriscau@unam.mx

III- Instituto Politécnico Nacional, San Pedro Zacatenco, México.
Contact: ferreiro@cinvestav.mx

La confiabilidad de la información en Internet: criterios declarados y utilizados por jóvenes estudiantes mexicanos¹

Marina Kriscautzky^{II}

Emilia Ferreiro^{III}

Resumen

La frecuencia del uso de Internet con propósitos de estudio entre jóvenes y niños pone de relieve la necesaria utilización de criterios de confiabilidad propios del medio. El tema es motivo de preocupación en el ámbito educativo y ha sido investigado, principalmente, con encuestas de opción múltiple. La originalidad del presente trabajo consiste en diseñar y aplicar un instrumento que permite un acercamiento a las discordancias entre datos de carácter declarativo y datos próximos a las decisiones a tomar en un contexto de acción. El TICómetro® es una encuesta en línea aplicada a 628 estudiantes mexicanos, entre 14 y 18 años de edad. Cuatro de las veintiséis preguntas del cuestionario son analizadas. Una de ellas remite a una lista de criterios para evaluar la confiabilidad de la información (opción múltiple); las otras tres preguntas exigen la selección de sitios Web confiables, con base en imágenes tomadas de Internet (screen shots). Es una situación más próxima a la búsqueda real en pantalla que permite contrastar los criterios declarados con los efectivamente utilizados. El análisis estadístico de los datos muestra que los criterios declarados no coinciden con los utilizados a la hora de seleccionar sitios Web confiables, en ambos grupos de edad. Sin embargo, aparecen diferencias significativas entre los dos grupos de edad en lo que concierne a los criterios enunciados. Otra contribución del presente trabajo consiste en proponer un nuevo agrupamiento de los datos declarativos, utilizando la distinción clásica entre texto y paratexto, proveniente de la teoría literaria, con los ajustes necesarios para tratar textos informativos en el espacio digital.

I- Este trabajo reporta parte de los resultados de la tesis de doctorado que la primera autora, desarrollada en el Departamento de Investigaciones Educativas del CINVESTAV bajo la dirección de la Dra. Emilia Ferreiro. Agradecemos a las autoridades de la Dirección General de Cómputo y de Tecnologías de Información y Comunicación de la UNAM y a todo el personal de la Coordinación h@bitat puma por hacer posible la aplicación del TICómetro®.

II- Universidad Nacional Autónoma de México, Ciudad de México, México.

Contacto: mkriscau@unam.mx

III- Instituto Politécnico Nacional, San Pedro Zacatenco, México.

Contacto: ferreiro@cinvestav.mx

Palabras clave

Confiabilidad de la información – Información digital – Internet – Jóvenes estudiantes.

Introduction

Using the Internet to find information for study has rapidly been incorporated into educational practices at different educational levels. At the university level, Internet searches have become indispensable for longer than a decade. But in compulsory education one also uses this technology with increasing frequency and at increasingly younger ages.

However, educators express concern about how young students perform such searches. They say that students merely place a search term in Google, enter the first result, copy and paste it, almost without reading what they have selected and without considering any questions about the validity of the information obtained. Do they do this because they are unable to address issues of validity and credibility? Does the way of posing the problem, in terms of teaching, contribute to make these issues invisible? There is educational research on children of the upper grades of primary and secondary education (COIRO, 2005; PERELMAN, et al., 2009; VEGA DÍAZ; ROJAS-DRUMMOND, 2012; COLWELL; HUNT-BARRON; REINKING, 2013) which shows the possibility of raising issues that promote group discussion and the implementation of strategies of search, selection and evaluation of the information they find. Children of those ages seem to be willing, under certain conditions, to consider the complex problematics concerning the credibility of information.

Greater autonomy to seek information online is expected from the young at higher educational levels. Therefore, it is important to know whether youth who have just completed secondary education and wish to enter college have resources to make decisions while selecting information for study. There are two ways to do this: surveys or some kind of interview. Surveys collect data from hundreds or thousands of individuals to describe a certain type of population. This quantitative advantage of surveys has the limitation of

obtaining declarative data through multiple choice questions. Of course, one can ask open questions but they are very difficult to analyze and categorize and thus be treated on the same level as closed questions. Individual interviews, in pairs or small groups, on the one hand, have the advantage of flexibility and adaptation to individual cases and, on the other, have the obvious disadvantage of not allowing the description of a numerically important group from a few cases.

We do not believe qualitative and quantitative methods to oppose in such a way that they are mutually exclusive. Instead, we have looked for ways to leverage the strengths of both. For that reason, in this research, the design of survey questions was preceded by some interviews to incorporate the point of view of young people into the questions. Not only that. The results of massive surveys also suggested the need for new interviews. We shall return to this. Data on the issue of credibility of information have been collected for more than a decade, especially in Europe and the United States. Most studies obtained data through multiple choice surveys, either on paper or online, with young college students (WATHEN; BURKELL, 2002; HAAS; WEARDEN, 2003; METZGER; FLANAGIN; ZWARUN, 2003; LIU, 2004; LIU; HUANG, 2005; BIDDIX; CHUNG; PARK, 2011; FRANCKE; SUNDIN; LIMBERG, 2011; HENKEL; MATTSON, 2011; FRANCKE; SUNDIN, 2012; EDWARDS, et al., 2013; LIM, 2013; LUCASSEN; SCHRAAGEN, 2013). As the age of Internet users has gone down so fast, other recent studies have focused on secondary students aged 11 to 18 years. (FLANAGIN; METZGER, 2011; TERRA; SÁ, 2012). PISA 2009 digital reading test, taken by students from Chile and Colombia (OECD, 2011), also provides data on the credibility of information among 15-year-olds.

In this article, we will report the results of an online survey of 628 students aged 14 to 18, aspiring to enter upper secondary education or university. In this sense, our

work is related to the previous investigations mentioned, particularly with regard to having a large sample of subjects surveyed. However, it differs in one key respect: the survey was designed with some questions which posed to respondents a task as close as possible to the actual selection of sites through images from webpages to be selected to obtain credible information for study. Both the selection criteria (in multiple choice questions) and the images of the proposed sites were defined from previously conducted qualitative interviews. We chose the context of study because accessing information to study is one of the reading purposes that imply the need to validate the credibility of information. When it comes to credibility, the declarative does not necessarily coincide with the operational. How do young people operate when they make decisions with a real screen? Do they use indicators which they somehow have learned to state?

What is meant by credibility?

The information available on the Internet consists of text, still and moving images and sound. However, in this work, we focus primarily on the search for textual information for study, as it occurs in almost all research articles on this topic.

What is meant by credibility of information in those articles? In English, which is the dominant language of the published works, the terms *credibility* or *believability* are used interchangeably and also relate to *trustworthiness and plausibility*. These terms are associated with a list of features that the information or message has to have to be credible or believable. In these lists, which are often very long, there are properties related to the source of the message, the receiver, the message itself, as well as to the opinions of authorities or peers, past experiences of the receiver with the site visited, time available to search and previous knowledge, beliefs or information. It is argued that believability

results from an “interaction between all these factors” (WATHEN; BURKELL, 2002, p. 140) and attempts to somehow classify so different elements are presented. For example, Tseng and Fogg (1999) group them around the four types of believability: presumed, reputed, surface and experienced. Terra and Sá (2012) also make a classification into four overlapping groups (conferred, voted, reputed and emergent believability).

Rieh and Beikin (1998, 2000) list seven criteria which *affect the believability* of information on the Internet, among which are the characteristics of the source, content, format and presentation, but they also include the speed of loading on the same level as the currency and accuracy of information.

How can one organize those lists of components of believability? Discussing the characteristics of the message and of the producer of that message seems inescapable, as well as talking about the characteristics of the receiver. However, in the list of characteristics of the latter, there are things such as cultural background and previous beliefs, indicating that it is very difficult to know what not to include.

In an attempt to offer a different approach to this way of describing the ingredients of credibility, we resorted to a completely different source and to a discipline seemingly unrelated to this discussion. The discipline is literary theory and the author is Gérard Genette (1982, [1987] 1997).

This author proposed the concept of *paratext* (1982 [1987] 2001), which is what surrounds the text itself and what is presented as such to readers, conditioning its reception because it provides readers with elements to anticipate text content. Referring to the literary work in print, as a book, Genette argues:

(...) this text is rarely presented in an unadorned state, unreinforced and unaccompanied by a certain number of verbal or other productions, such as an author's name, a title, a preface, illustrations.

And although we do not always know whether these productions are to be regarded as belonging to the text, in any case they surround it and extend it, precisely in order to *present* it, in the usual sense of this verb but also in the strongest sense: to *make present*, to ensure the text's presence in the world, its "reception" and consumption in the form (nowadays, at least) of a book. These accompanying productions, which vary in extent and appearance, constitute what I have called elsewhere the work's *paratext*, in keeping with the sometimes ambiguous meaning of this prefix in French. (GENETTE, [1987] 1997, p.3)

Alvarado (1994) returns to Genette and defines the paratext as a reading guide, i.e., the set of elements that anticipate the content of the text:

Threshold of the text, the reader's first contact with the printed material, the paratext is instructive, it is a reading guide. (...) This is particularly evident in the case of print (...) but the literary, scientific or dissemination genres also provide readers, from its format, with recognition elements and the opportunity to make first hypotheses about the content of the text, which the reading will confirm or refute a posteriori. (ALVARADO, 1994, p.2)

In the case of printed material (books, magazines, journals, etc.) paratextual elements relate to the characteristics of the graphic format that define each text type, graphic elements (images, pictures) that accompany the text, marks that identify the publisher and author, the publication date, together with all those texts (prefaces, reviews, comments) that relate to the work and provide readers with elements to interpret it. In the digital text, some paratextual elements of the object book are preserved, but new elements appear. One of them, of particular interest in the context of this research, is the

address or URL, which allows anticipating what kind of website it is: whether it has a supporting institution, which country it comes from and even whether it is a reputable site, or conversely, a site that does not deserve to be referred to.

Paratextual elements have an anticipatory nature with regard to the content of the text because they allow readers to formulate hypotheses regarding the information, in terms of both its content and quality. In the latter sense, the paratext can be considered important when it comes to evaluating the credibility of information.

The distinction between textual and paratextual elements coincides in some way with the perspective from which the evaluation of the credibility of information has been addressed in various research fields (psychology, communication, library science, education), distinguishing at least two basic elements: the credibility of the message (quality of information, writing, type of argument) and the credibility of the source (author, editor) (HOVLAND; JANIS; KELLEY, 1953). Other authors introduce the medium as a third element to be evaluated. (WATHEN; BURKELL, 2002; FLANAGIN; METZGER, 2007; HILLIGOSS; RIEH, 2008). By *medium* they mean the ways in which information flows: TV, Internet, print. It might seem, then, that we can distinguish between the credibility of the message itself (textual elements) and the credibility of the medium and the source (paratextual elements that allow anticipating the credibility of information). The set of paratextual elements is not easy to establish on the Internet, but we believe that the above distinction contributes to the necessary theoretical discussion on the ways to conceptualize credibility.

Consequently, it seems appropriate to maintain the distinction between the criteria for evaluating the credibility of the text, message or content and the criteria for evaluating the credibility related to the paratext.

In the design of the survey questions, we considered these two aspects (see *Methodology*,

where we describe in these terms the criteria in question 14):

a) assessment of textual elements (how it is written, how understandable it is for the reader, how it responds to the information needs of the seeker, writing in the language of the country where the survey is conducted or in English, the dominant language of the international scientific community);

b) assessment of the paratextual elements: what type of site it is (educational, research, encyclopedia, or blog, magazine, commercial); identification data (author, creator of the site, supporting institution); date of publication; presence or absence of advertising, images or multimedia elements, along with typographic features such as color, size, type of font.

Methodology

We designed an online survey called TICómetro®. To develop it, we used the Moodle platform. This is an open-source platform that has a module to develop questionnaires in which the order of the questions and answer choices (images or text) changes each time the tool is opened. Therefore, the distribution of the answer choices is different for each respondent. This procedure of random ordering distinguishes this study from all the studies previous to this one (FLANAGIN; METZGER, 2011; BRATEN; STROMSO; SALMERÓN, 2011; KUBISZEWSKI; NOORDEWIER; COSTANZA, 2011; TERRA; SÁ, 2012).

To enable the display of images from websites, we developed a script that allows the integration of images as answer choices and their expansion when clicking on each one.

TICómetro® was answered by 302 young people aspiring to enter Universidad Nacional Autónoma de México (UNAM) at high school level (14 years) and by 326 at undergraduate

level (18 years), a few weeks before the start of the school year. In total, 628 young people aged from 14 to 18 years. 86% of the surveyed population comes from free public schools. The young voluntarily participated in a stand of the Educational Guidance Fair organized annually by UNAM to inform young students about its educational provision. In such stand, we proposed to know their skills in using information and communication technologies. They automatically obtained a score expressed in ranges: beginner, intermediate and expert. There were four personal computers and someone available to help each respondent, if necessary. There was no time limit for responding. On average, each participant took from 15 to 20 minutes.

The tool consists of four items of general information: gender, age, educational level they aspire to (high school or college) and type of school of origin (public or private). These items appeared at the beginning and in the same order. Students then answered 26 questions related to four themes: online communication and collaboration; information processing; safe use of computers and the Internet; access, selection and evaluation of information. The 26 questions were in a different order each time a student entered the questionnaire. Not only the questions but also the answer choices for each question appeared in a different order each time.

In this article only questions 11, 14, 16 and 18 will be analyzed because they are the ones which focus on the research question: the criteria for evaluating the credibility of Internet sites for study purposes.

Question 14 (*To select credible information, you look at...*) reproduces the model of declarative information with multiple choice answers. Respondents were asked to select one to three options from a list of criteria for assessing credibility.

Answer choices of question 14.

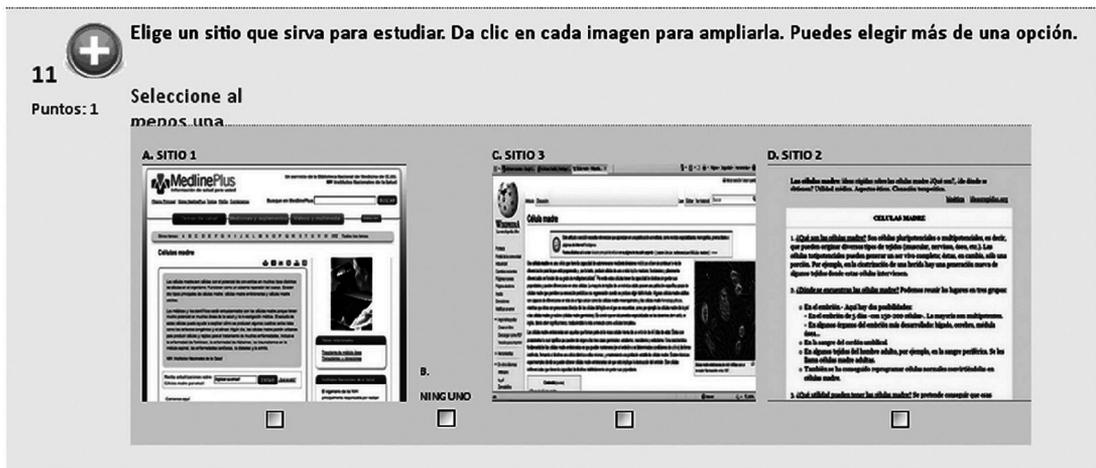
- a. If the URL or address is .edu, .gov, or a university.
- b. If it brings all the information I'm looking for.
- c. If the site has photos or images.
- d. If the information it brings is easy for me to understand.
- e. If the website design is attractive.
- f. If it has an author.
- g. If the information has been published recently.
- h. If the information is well written.

The option list was built based on the criteria reported in the literature analyzed, data from previous qualitative interviews and the distinction between textual and paratextual elements. For expository purposes, we showed the options ordered according to the latter criterion:

- If the information it brings is easy for me to understand (textual).
- If the information is well written (textual).
- If it brings all the information I'm looking for (textual).
- If the design is attractive (paratextual).
- If the information has been published recently (paratextual).
- If the URL or address is .edu, .gov or a university (paratextual).
- If the site has photos or images (paratextual).
- If it has an author (paratextual).

Questions 11, 16 and 18 are the most original ones of the tool. These are trios of images from websites and the respondent must choose one or two. Question 11 asks respondents to choose sites to study, without explicit reference to credibility.

Figure 1. Choose a website to study. (Question 11).



Images of three websites to select. Clicking on each image enlarges it to a size of 800 x 600 pixels. Source: Authors, based on the 2011 version of TICómetro®.

The websites presented in this question are Wikipedia, MedlinePlus and *Ideas rápidas*. The fourth alternative is the option *none*. All visible pages refer to the same content: stem cells, a theme which is not part of the Biology content of compulsory education yet, but that has had a growing social presence for

its medical applications. We chose Wikipedia because it was the most accessed site among the respondents in our previous research and because is a site that provokes debate about credibility (KUBISZEWSKI, NOORDEWIER y COSTANZA, 2011; FRANCKE y SUNDIN, 2012; LIM, 2013; LUCASSEN y SCHRAAGEN, 2013).

MedlinePlus (Figure 2) is a virtual encyclopedia of medical subjects published by the U.S. National Library of Medicine, and renowned in the health field for its updated and accurate information.

Therefore, there was a contrast between a popular site (Wikipedia), a credible site for a scientific community (MedlinePlus) and a site that does not meet the criteria of paratextual credibility, but offers understandable and

Figure 2



Image from MedlinePlus site.

Source: screenshot from: <www.nlm.nih.gov/medlineplus/spanish/stemcells.html> (Accessed: September 2011)

Figure 3

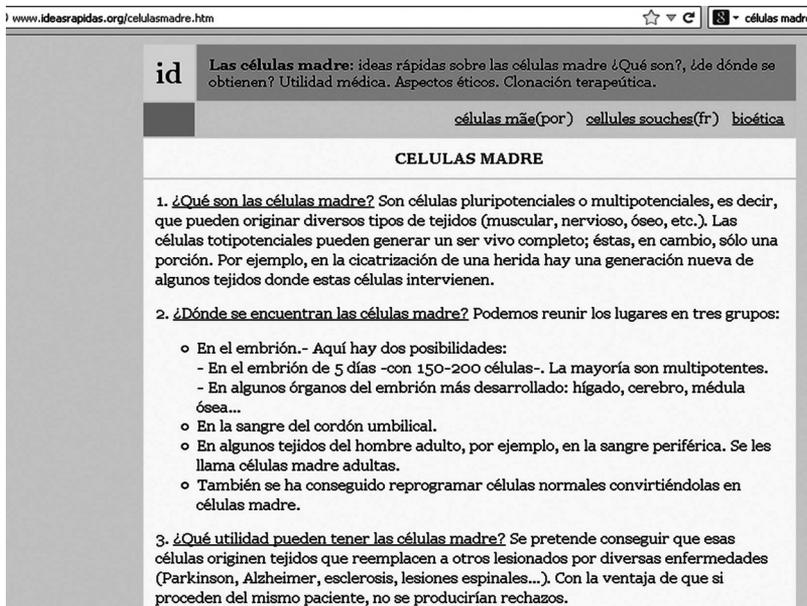


Image from ideasrapidas.org website (Accessed: September 2011) Source: <screenshot of http://www.ideasrapidas.org/celulasmadre.htm>

practical textual information. Questions 16 and 18 explicitly address the selection of credible information regarding the same content: the mummies of ancient Egypt. This content is not only present in the compulsory education curriculum but is also very popular among children and youth. Additionally, one should keep in mind that in Mexico there are natural mummies, found in Guanajuato state,

which receive many Mexican and foreign visitors.

Question 16 is formulated as follows: *In what site do you think you can find credible information?* It has three options of website images. One is of an Egyptian periodical published in English, which reports the discovery of a mummy and includes three photos. (Figure 4).

Figure 4

The screenshot shows the Al-Ahram website interface. At the top, there is a logo for Al-Ahram Weekly On-Line, the date 17-23 August 2006, Issue No. 808, and the Heritage category. A search bar and a 'Google Search' button are also present. Below the header, there are links for 'Recommend', 'Comment', and 'Printer-friendly'. The main content area features an article titled 'The elite of the millennia' with a sub-headline: 'The second Royal Mummies Hall which opened last week crosses the Ts on mummification and reveals an obscure era of instability in the history of ancient Egyptians. Nevine El-Aref toured the gallery'. The article text discusses the importance of mummification in ancient Egypt and mentions the New Kingdom. To the right of the text are three images: a mummy, a tomb interior, and a mummy in a glass case. A vertical navigation menu on the right side lists various categories: Front Page, Egypt, Region, Lebanon, Economy, Opinion, Reader's corner, Press review, Culture, Features, Heritage, Living, Sports, Cartoons, Encounter, People, Listings, and BOOKS.

Image from the website of Egyptian newspaper Al-Ahram.
Source: screenshot of <<http://weekly.ahram.org.eg/2006/808/hr2.htm>>; (Accessed: September 2011)

The second one is of the British Museum, internationally renowned, in the section of their official site dedicated to Egyptian mummies, with very brief information in English and photos. (Figure 5).

Both sites are in English because we wished to investigate what aspects respondents assess when they cannot easily understand the content of the text because

it is in a foreign language. In this sense, these two sites make it considerably difficult to assess the textual aspects. Finally, the last option is the popular Yahoo Answers site (Figure 6), also with a question about Egyptian mummies. It is well known that on this site anyone can post a question and the general public answers based on their experience.

Figure 5



Image from The British Museum site.

Source: screenshot of <http://www.britishmuseum.org/explore/young_explorers/childrens_online_tours/journey_into_the_mummy/journey_into_the_mummy.aspx>. (Accessed: September 2011)

Figure 6



Image from Yahoo Answer website.

Source: <screenshot of <http://mx.answers.yahoo.com/question/index?qid=20061208132815AA6EbA0>>. (Accessed: September 2011)

Figure 7

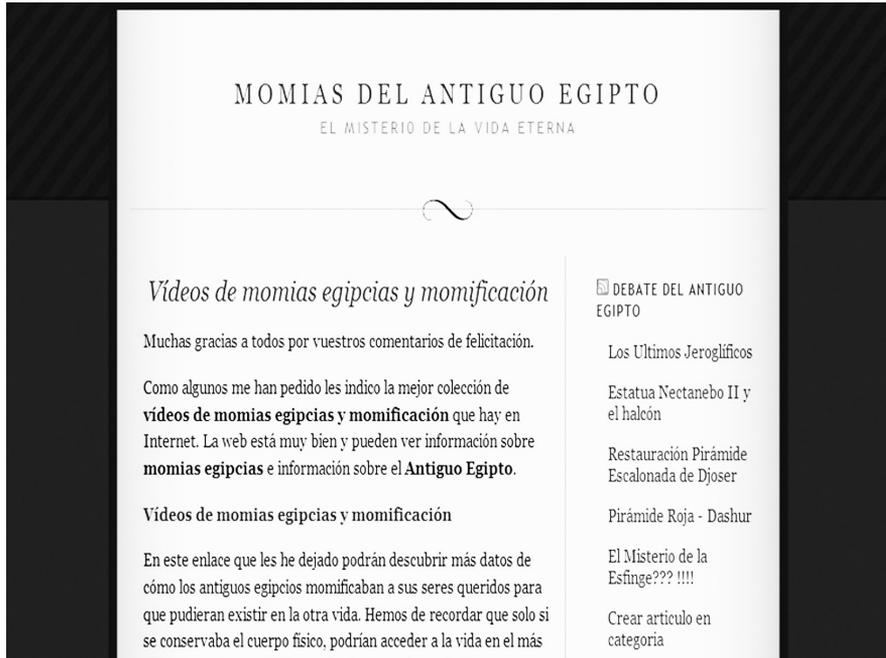


Image from the blog <http://momia.wordpress.com/>
Source: <screenshot of <http://momia.wordpress.com/>>. (Accessed: September 2011)

Figure 8



Image from monografias.com site
<<http://www.monografias.com/trabajos42/momificacion-egipto/momificacion-egipto.shtml>> (Accessed: September 2011).

Question 18 repeats the task of selecting credible information from images with another group of webpages: a blog (Figure 7), Yahoo Answers (the same as in the previous question) and monografias.com (Figure 8), a commercial site which posts homework that users can download.

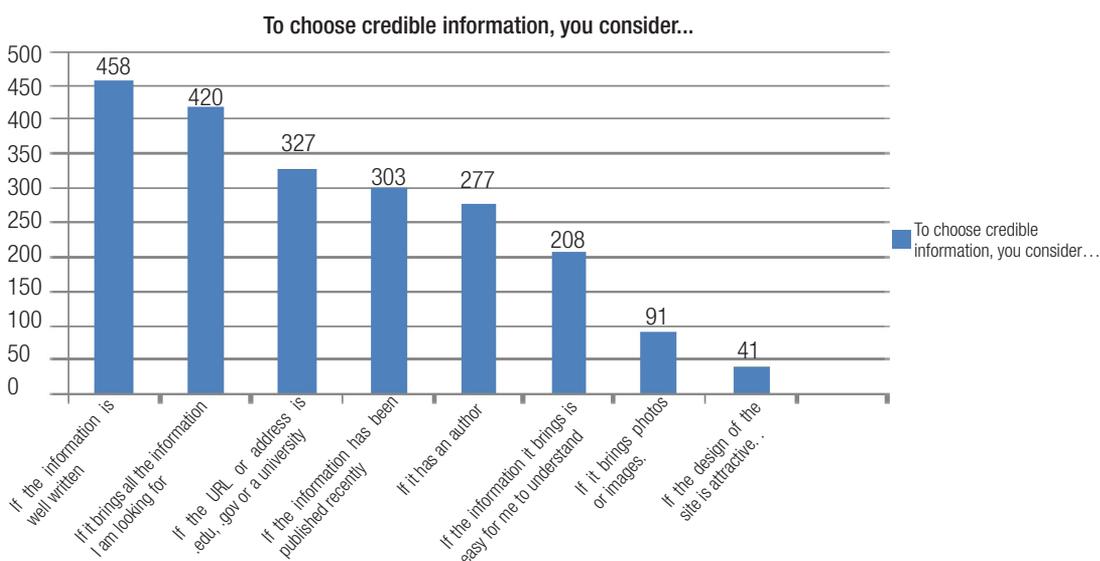
The three sites were presented under the assumption that respondents would choose the option *none* because they lack elements to provide believability. All three are in Spanish, which allows evaluating textual elements. Regarding paratextual data, in the three options, there are elements to assess the poor credibility of the sites and, therefore, of the information they contain: the URL, the lack of data on the author and/or institution, the absence of publication dates and the characteristics of their graphic formats.

We expected that the contrast between the declarative responses and those that approximate the behavior of a real decision-making context would bring us closer to the answer we were seeking.

Results

Chart 1 shows the criteria chosen most frequently by all respondents for selecting credible information in a context of multiple choice between statements. Good writing appears as the most selected criterion (*option: the information is well written*). The following criterion in order of frequency also has a textual nature: information is credible if it meets the needs of the seeker (*option: it brings all the information I'm looking for*). Such criterion is followed by the consideration of the URL, publication date, and the location or identification of the author, which are all paratextual elements. The publication date is possibly linked to the knowledge that young people have about the instability of Internet publications or their daily realization that information is rapidly produced and changes rapidly too. Locating or identifying the author might be a consequence of the experience that anyone can write on the Internet and of the criteria that prevail in the world of printed texts, in which the author is an important datum to decide on the credibility of the information provided.

Chart 1. Number of times the 628 respondents selected each of the options in the list of declarative criteria for selecting credible information. (Question 14)



Source: Authors, based on the 2011 version of TICómetro ®.

Data are presented in order of frequency, in absolute numbers. They are not additive because each respondent chose from one to three options. See Appendix 1, which shows one of the combinations that were presented at random to students.

Only 413 out of 628 respondents selected three options. With this subset we analyzed the most frequently selected combinations. The fact that respondents could choose three out of eight options generated 56 possible combinations. If each combination had been chosen with equal frequency, we would have had 7.37 students for each combination. In our sample, 49, 43 and 41 students selected the three most frequent combinations. These frequencies are relevant for being seven times higher than the expected number under the scheme of equiprobability in the selections in each triple.

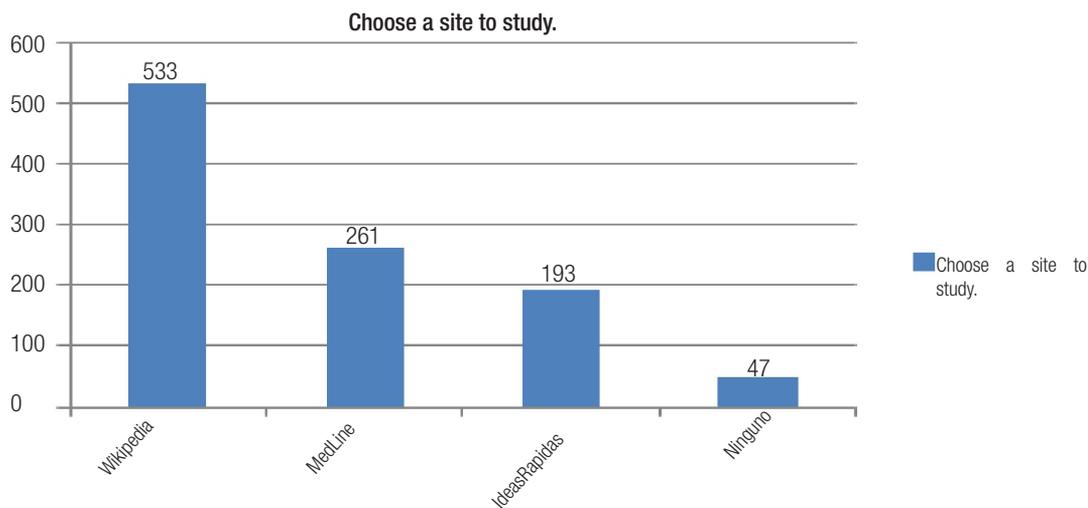
The first combination, in order of frequency, is: *the information is well written + it brings all the information I'm looking for + the information that it brings is easy for me to understand*. These three textual criteria can only be applied while the information is read. Such combination is followed by the combination

of the following options: *the information is well written + it has an author + the URL or address is .edu, .gov, or a university...* Here, two paratextual criteria are combined with a textual one. The third combination of answers is: *the information is well written + it brings all the information I'm looking for + the information has been published recently*. Such combination contains two textual criteria and one paratextual criterion.

The three most frequent combinations include a textual criterion linked to the quality of writing. Although we cannot infer with certainty what the youth surveyed understood by *well-written information*, there is no doubt that they mean specifically the textual aspects. Additionally, two of the most frequent combinations incorporate another textual element, which includes the reader's expectations with respect to information (*it brings all the information I'm looking for*).

However, when respondents were faced with the task of selecting a site to study from website images, the aforementioned criteria were not always taken into account, as shown in Chart 2.

Chart 2. Number of times the 628 respondents selected each of the three options of images from websites to study. (Question 11)



Source: Authors, based on the 2011 version of TICómetro®.

Data are presented in order of frequency, in absolute numbers. They are not additive because each respondent selected one or two options, or the option *none*.

Chart 2 shows that the majority of respondents (85%) select Wikipedia. The difference in the frequency of choice of the other two sites is marked. There is little difference between choosing MedlinePlus and ideasrapidas.org although one site is very different from the other in all its paratextual and textual elements. *Ideas rápidas*, as shown in Figure 3, does not contain data of the site, its author/creator and it does not have a publication date or clues of a supporting institution.

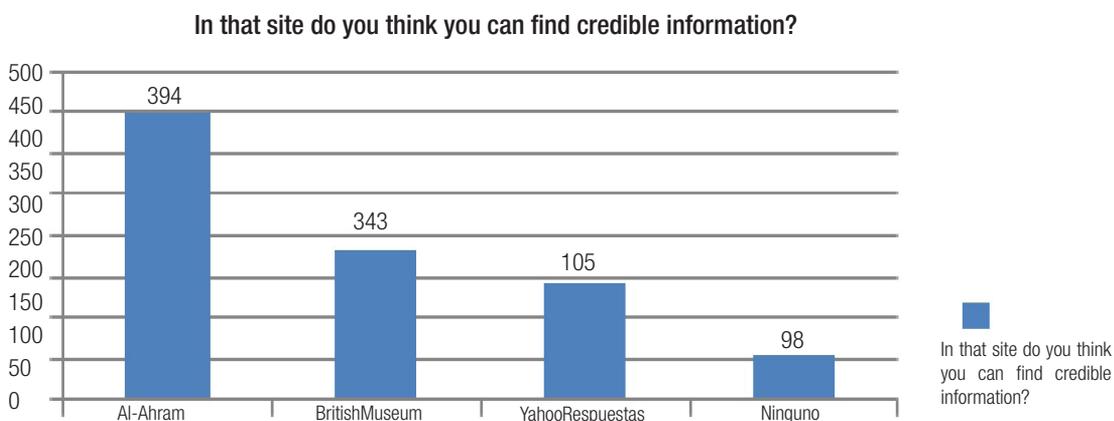
By contrast, as shown in Figure 2, MedlinePlus is a site where one finds paratextual elements which support the credibility of information: it is a .gov site, it

is clearly identified as a service of the U.S. National Library of Medicine, it has a link to the section *About MedlinePlus*, where one can find out who makes the site and who is responsible for the information, as well as a contact section, which allows readers to communicate with the developers of the site.

The little difference between the selection of MedlinePlus or ideasrapidas.org leads us to think that, in a concrete situation of selection on the screen, paratextual declarative criteria do not appear.

Question 11 speaks of *sites for studying* without explicitly mentioning credibility. In contrast, questions 16 and 18 explicitly mention credibility and allow us to see whether there is agreement or discrepancy between the criteria stated and the criteria used to choose sites with credible information. Figures 3 and 4 present the data.

Chart 3. Number of times the 628 respondents selected each of the three options of images from websites where they believe they can obtain credible information. (Question 16)



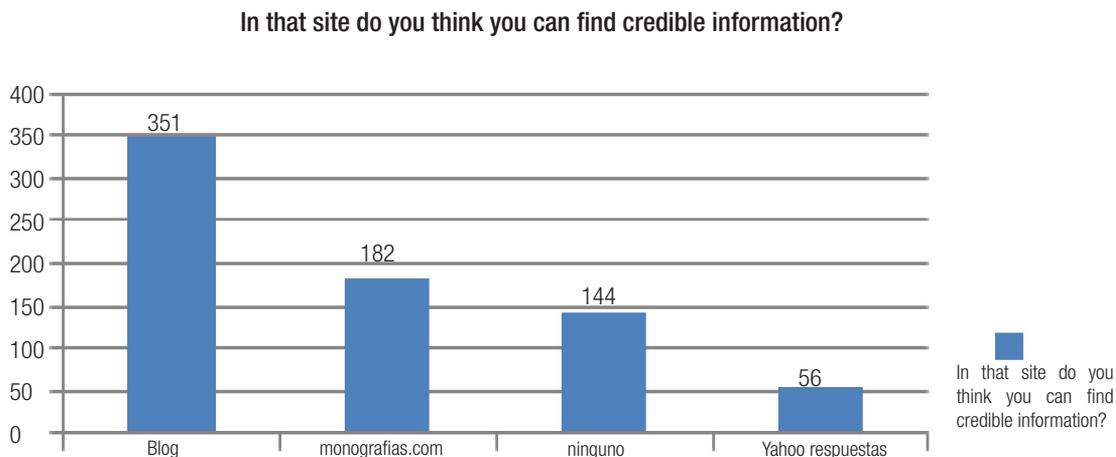
Source: Authors, based on the 2011 version of TICómetro®.

Data are presented in order of frequency, in absolute numbers. They are not additive because each respondent chose one or two options, or the option *none*.

Interestingly enough, the two most chosen websites are those of the Egyptian periodical and the British Museum, both with

images on the same topic. Both are in English, which makes it very difficult to access textual content. On the site of the periodical, one can see the publication date and other graphic elements that indicate that it is an online periodical, where one can find news (see Figure 4). The name of the museum is in English but the word

Chart 4. Number of times the 628 respondents selected each of the three options of images from websites where they believe they can obtain credible information. (Question 18).



Source: Authors, based on the 2011 version of TICómetro®.

is almost the same in Spanish; photographs show objects of a museum exhibition and the design is consistent with an institutional site. (See Figure 5).

Data are presented in order of frequency, in absolute numbers. They are not additive because each respondent chose one or two options, or the option *none*.

Finally, in question 18, we include three sites that do not meet the criteria of credibility, under the hypothesis that they could choose the option *none*. Even though this option was chosen 144 times, it is far from being the majority choice. The blog was chosen 351 times. The blog page (like all the others) mentions the same content: the mummies of ancient Egypt. It was designed in one of the most used services for creating blogs (Wordpress) in a very common template, which gives it the appearance of a book. The title and subtitle could be of a book, but the text is written in first person, in the dialogic form (it starts with: *Thank you all for your comments*). Only the similarity with the page of a book and the title (Ancient Egyptian Mummies) could justify the choice, as none of the criteria stated are reflected in this blog. (See Figure 7).

With respect to monografias.com, the distance between the criteria stated and the characteristics of this site is enormous. Although it was selected by 182 subjects, the truth is that it should have been chosen by nobody because: it is surrounded by advertisements, including an offer to participate in a contest and earn dollars; it does not contain textual information despite its title being similar to the previous one. Therefore, none of the stated criteria may have been used (see Figure 8). Something similar can be said about Yahoo Answers, which was elected 56 times.

We expected question 18 to have a considerable number of responses rejecting the options offered. We had just 144. This could be due to the fact that respondents interpreted this question as *I have to choose one of the options provided*. However, we can also interpret this result, together with the previous ones, as a confirmation of the main hypothesis: the criteria young people stated to be necessary to select credible information do not come into play every time they have to choose a site.

Analysis of the data by type of school of origin and gender did not show significant differences. However, there are differences

that deserve to be taken into account when we compare the responses of the 14-year-old students, applying for high school, to those of the young aged 18, applying for college.

14 out of the 413 students who chose three possible answers in question 14 on criteria for knowing when the information is credible, 210 are prospective college students and 203 are prospective upper secondary students, i.e., high school students. With the totals for each subset, we calculated the absolute and relative frequencies of selection of the three most chosen combinations. We observed that the most chosen combinations are different in each age group. Among the 18 year-olds, the most chosen combination is: *The information is well written + it has an author + the URL or address is .edu, .gov, or a university.* The second combination in order of frequency is: *the information has been published recently + it has an author + the URL or address is .edu, .gov, or a university..* The third is *the URL or address is .edu, .gov, or a university + the information is well written + it brings all the information I'm looking for...*

In each of these three combinations, one paratextual criterion appears: website analysis based on the URL or address, which belongs exclusively to the digital field. In one of the combinations, that criterion appears with two other paratextual criteria: publication date and author. In the other two combinations, the reference to the URL includes a textual criterion: good writing.

Table 1. Most frequent combinations of criteria of credibility among prospective college students.

Prospective college students	Absolute frequency	Relative frequency
Writing, author, URL,	30	0.14
Date, author, URL	23	0.11
URL, writing, comprehensiveness	21	0.10

Absolute and relative frequencies of selection of combinations of criteria of credibility among prospective college students. Total = 210
Source: Authors, based on the 2011 version of TICómetro®.

Let us now consider the differences with the young aged 14, applying for upper secondary education (high school). In this case, the combinations are:

Table 2. Most frequent combinations of criteria of credibility among high school students.

High school students	Absolute frequency	Relative frequency
Writing, understandable comprehensiveness	30	0.15
Writing, comprehensiveness, date	26	0.13
Writing, author, URL	13	0.06

Absolute and relative frequencies of selection of combinations of criteria of credibility among high school students Total = 203

Source: Authors, based on the 2011 version of TICómetro®.

In this group, all combinations include textual criteria. In the three combinations, the good writing of information is present. The paratextual criterion, the URL analysis, dominant in the group aged 18, appears in only one of the combinations, the least frequent one.

Consequently, older subjects tend to consider paratextual aspects and the younger ones focus more on textual aspects. This could indicate that the older ones can, at least from the declarative point of view, consider elements predictive of the credibility of information, such as the URL. The younger ones focus on elements of the text that demonstrate a concern for achieving the reading purposes (the information is understandable, is comprehensive and well written) rather than for evaluating credibility.

It remains for us to consider the agreement or disagreement between saying and doing, i.e., between the criteria that they state they take into account and the criteria that they actually use to select a credible site.

To verify this, we conducted an analysis of independence between two variables using an expression for the stochastic independence $P(A \cap B) = P(A) * P(B)$, where P is the probability of selecting a criterion or an

option, A is the set of triples of declarative criteria and B is the set of options of images of websites. This was done in order to determine the relation between the criteria stated in question 14 and site selection in question 18. This question, as we have already mentioned, was designed under the assumption that they would choose the option *none*, as the sites proposed do not meet the criteria for

evaluating credibility which were reported in the literature and included in the options in question 14, which are of declarative type. Table 3 shows, for the group of prospective college students, the calculation of the probability of selection of each of the three combinations of declarative criteria and the probability of selection of each of the options in question 18.

Table 3. Relative frequencies of selection of the four options in question 18.

Prospective college students	Blog	Monografias.com	Yahoo	None	Totals
Writing, author, URL	0.2973	0.0676	0.0000	0.0405	0.4054
Date, author, URL	0.1081	0.0811	0.0135	0.1081	0.3108
URL, writing, comprehensiveness	0.1216	0.0405	0.0541	0.0676	0.2838
Totals	0.5270	0.1892	0.0676	0.2162	1.0000

Calculation of the probability of selection of each triple of credibility criteria by the probability of selection of each response option for the group of prospective college students. (Total = 210)

Source: Authors, based on the 2011 version of TICómetro®.

With these data, we verified the ratio between each of these combinations and the option *none*, this being the best answer under the criteria for selecting the sites presented. This is what Table 4 shows.

Table 4. Independence relation between declarative criteria and the option *none* in question 18.

Prospective college students	P None	P Triple	P (None= 0.2162)*P(Triple)
Author, URL, writing	0.0405	0.4054	0.0877
Date, author, URL	0.1081	0.3108	0.0672
URL, writing, comprehensiveness	0.0676	0.2838	0.0614
Totals	0.2162	1.0000	

Probability of selection of each triple of credibility criteria by probability of selection of the option *none* in question 18 for the prospective college group. (Total = 210)

Source: Authors, based on the 2011 version of TICómetro®.

Table 4 shows the ratio between the declarative criteria and the selection of the option *none*. In the first two triples or combinations (0.0405 – 0.0877; 0.1081 to 0.0672), figures are far from similar, indicating that taking into account the criteria stated in the first two combinations influences the probability of selecting the option *none*. In contrast, in the case of the third triple, results are similar (.0676 to .0614). This could indicate that there is a relation of independence between this triad of criteria and the selection of the option *none*.

Let us see what happens in the younger group, those who are high school students. Table 5 shows the calculation of the probability of selecting each of the three combinations of declarative criteria in relation to the probability of selecting each of the options in question 18.

Table 5. Relative frequencies of selection of the four options in question 18.

High school students	Blog	Monografias.com	Yahoo	None	Totals
Understandable, writing, comprehensiveness	0.1159	0.1594	0.1014	0.0580	0.4348
Writing, comprehensiveness, date	0.1739	0.0725	0.0435	0.0870	0.3768
Writing, author, URL	0.0580	0.0145	0.0145	0.1014	0.1884
Totals	0.3478	0.2464	0.1594	0.2464	1.0000

Calculation of the probability of selection of each triple of credibility criteria by the probability of selection of each response option for the high school group. (Total = 203)

Source: Authors, based on the 2011 version of TICómetro®.

With these data, we proceeded to verify the relation between each of these combinations and the answer *none*. This is shown in Table 6.

Table 6. Independence relation between declarative criteria and the option none in question 18.

High school students	None	Totals	P (None= 0.2464)*P (Triple)
Understandable, writing, comprehensiveness	0.0580	0.4348	0.1071
Date, writing, comprehensiveness	0.0870	0.3768	0.0928
Author, URL, writing	0.1014	0.1884	0.0464
Totals	0.2464		

Probability of selection of each triple of credibility criteria by the probability of selection of the option *none* in question 18 for the high school group.

Source: Authors, based on the 2011 version of TICómetro®.

In the younger group, the results are different in the three triples. Therefore, there is no consistency between the declarative criteria and the probability of selecting the option *none*, since any of the triples of declarative criteria may influence the selection of option *none*. We can confirm the hypothesis that what they may declare is different from what they may consider when actually selecting credible sites.

Conclusions

The data show an interesting difference between the two age groups (14 and 18),

and such difference does not appear in other published works (e.g., FLANAGIN and METZGER, 2011). The younger ones refer to the characteristics that textual information must have to be accessible to them as readers: being well written, easy to understand and providing comprehensive information. These criteria are typical of textual information, both on the screen and on paper. A specific property of the digital medium that allows us to anticipate credibility before starting to read (URL) is taken into account preponderantly by the older subjects. This could be due to many factors which we can only speculate about. It could

be that these young people have accumulated more time browsing the Internet, that its use for study has been fostered during their years of secondary education, that considering the specific properties of the digital medium has been effective on previous occasions to save time when seeking information.

However, the discrepancy between what the young state and the selection they make in a context closer to that of the real search is common to both age groups. This suggests that it is not enough to tell students to find information on the Internet. It would be desirable to devise teaching activities for students to be faced with problems that allow them to discuss the issue of credibility and implement what they state. In such an unstable and changing environment as the Internet, where the supply of webpages and modes of interaction are constantly growing and changing, one cannot imagine a closed list of criteria to consider. Rather, we must learn to identify properties of digital information in order to decide, now and in the unpredictable future, whether such information is credible.

This lack of consistency between what one may state and what is actually done in concrete situations should be taken into account in research, because it calls into question the way in which data are obtained, often using what students state and not what they actually do. Of course, we have not observed the youth doing a real search with specific purposes, but, in a context of mass survey, the questions analyzed in this study are the closest we managed to get to that. Preliminary data from the interviews that we are carrying out with younger children, from primary education, hint at a discrepancy between declarative criteria which, in many cases, are the simple repetition of verbal information gathered from other adults or peers (for example, Wikipedia cannot be trusted because anyone can write there) and the criteria they implement when they conduct an actual search and choose Wikipedia, because it always appears among the first results and

they never go beyond the first page of results. Finding these discrepancies is one thing; trying to understand how they manage to justify what they finally decide is something else. This and other problems of qualitative nature require semi-structured interviews with a small number of subjects, preferably in pairs to foster the exchange of views.

We believe that our work provides important insight into research on the credibility of information on the Internet. First, the construction of the research tool allowed us to provide respondents with a situation very close to the actual selection of credible sites through the presentation of screenshots of websites. The tool devised, TICómetro®, allowed us to collect data with relatively large samples, which could not be obtained in cases of observation of individuals searching information on the Internet.

This tool does not replace observation and interviews with young people in their actual search for information and its assessment. In fact, we are conducting such interviews not only with young people but also with children aged 9 to 12 years, since the age of Internet users has been decreasing sharply.

To try to put into order the lists of elements or factors that influence the perception or judgment of believability, we have resorted to the distinction between textual and paratextual, from another disciplinary field (literary theory) and from a time previous to information technology. In the analysis of the responses, we have used the terms above, with the necessary adaptations to informational texts and to the digital medium. It is an issue that undoubtedly requires further development, but a first approximation has allowed us to find significant links between certain answers, without losing the necessary distinctions.

There remain several research tasks. We shall mention just two of them. On the one hand, to reduce the age of the subjects interviewed, adapting the methodology to different age ranges. On the other hand, most research focuses on written information (with a few exceptions,

such as PERELMAN et al, 2009). Despite methodological difficulties, it is necessary to pay attention to the large amount of information circulating on the Internet in the form of images (still photos, videos, drawings and hybrids). These images also convey information for study purposes, and the credibility criteria used by students need to be studied.

The data presented are relevant also with regard to education. Evaluating the credibility of information is a challenge for students, even

at higher education level. In particular, it is a challenge difficult to overcome in action. Even when they have criteria for selecting credible information on the declarative level, when in action these criteria can compete with practical requirements or unique motivations. Clearly, it is necessary to design teaching activities that promote the development of one of the tasks of the current reader: to distinguish what information is credible in contexts of Internet search with several purposes, particularly with study ones.

References

ALVARADO, Maite. **Paratexto**. Buenos Aires: Universidad de Buenos Aires, Instituto de Lingüística, Facultad de Filosofía y Letras, 1994.

BIDDIX, Patrick; CHUNG, Joo; PARK, Han Woo. Convenience or credibility? A study of college student online research behaviors. **Internet and Higher Education**, v. 14, n. 3, p. 175-182, 2011. Disponible en: <<http://www.sciencedirect.com/science/article/pii/S1096751611000042>>. Acceso en jul. 2012.

BRATEN, Ivar; STROMSO, Helge; SALMERÓN, Ladislao. Trust and mistrust when students read multiple information sources about climate change. **Learning and Instruction**, v. 21, n. 2, p.180-192, 2011. Disponible en: <<http://www.sciencedirect.com/science/article/pii/S0959475210000162>>. Acceso en jul. 2012.

COIRO, Julie. Making sense of online text. **Educational Leadership**. Association for Supervision and Curriculum Development, v. 63, n. 2, p. 30- 36. oct. 2005. Disponible en: <<http://www.ascd.org/publications/educational-leadership/oct05/vol63/num02/Making-Sense-of-Online-Text.aspx>>. Acceso en mar. 2009.

COLWELL, Jamie; HUNT-BARRON, Sarah; REINKING, David. Obstacles to developing digital literacy on the internet in middle school science instruction. **Journal of Literacy Research**, v. 45, n. 3, p. 295-324, 2013. Disponible en: <<http://jlr.sagepub.com/content/45/3/295.abstract?rss=1> >. Acceso en: nov. 2013.

EDWARDS, Chad et al. How much klout do you have... a test of system generated cues on source credibility. **Computers in Human Behavior**, v. 9, n. 5, p. A12-A16, 2013. Disponible en: <<http://www.sciencedirect.com/science/article/pii/S0747563212003767>>. Acceso en: oct. 2013.

FLANAGIN, Andrew J.; METZGER, Miriam. The role of site features, user attributes, and information verification behaviors on the perceived credibility of web-based information. **New Media & Society**, v. 9, n. 2, p. 319-342, 2007. Disponible en: <<http://www.comm.ucsb.edu/faculty/flanagin/CV/FlanaginandMetzger2007%28NMS%29.pdf>>. Acceso en: mar. 2009.

FLANAGIN, Andrew J.; METZGER, Miriam. Kids and credibility: an empirical examination of youth, digital media use, and information credibility. In: FLANAGIN, Andrew J. et al. **The John D. and Catherine T. MacArthur Foundation Reports on Digital Media and Learning**, Cambridge: MIT Press, 2011. 135 p.

FRANCKE, Helena; SUNDIN, Olof. Negotiating the role of sources: educators' conceptions of credibility in participatory media. **Library & Information Science Research**, v. 34, n. 3, p. 169-175, 2012. Disponible en: <<http://lup.lub.lu.se/luur/download?func=downloadFile&recordId=2371826&fileId=2596842>>. Acceso en: jul. 2012.

FRANCKE, Helena; LIMBERG, Louise. Debating credibility: the shaping of information literacies in upper secondary school. **Journal of Documentation**, v. 67, n. 4, p. 675-694, 2011. Disponible en: <<http://www.emeraldinsight.com/journals.htm?articleid=1941437>>. Acceso en: ene. 2012.

GENETTE, Gerárd. **Palimpsestes**: la littérature au second degré. Paris: Éditions du Seuil, 1982.

GENETTE, Gerárd. **Umbrales**. 1. ed. México: Siglo XXI, 2001.

HAAS, Christina; WEARDEN, Stan. E-credibility: building common ground in web environments. **L1- Educational Studies in Language and Literature**, Netherlands, Kluwer Academic Publishers, n. 3, p. 169-184, 2003.

HENKEL, Linda; MATTSON, Mark. Reading is believing: the truth effect and source credibility. **Consciousness and Cognition**, v. 20, n. 4, p. 1705-1721, 2011. Disponible en: <<http://www.sciencedirect.com/science/article/pii/S1053810011002200>>. Acceso en: ene. 2012.

HILLIGOSS, Brian; RIEH, Soo Young. Developing a unifying framework of credibility assessment: construct, heuristics, and interaction in context. **Information Processing and Management**, v. 44, n. 4, p. 1467-1484, 2008. Disponible en: <http://rieh.people.si.umich.edu/papers/hilligoss_ipm.pdf>. Acceso en mar. 2009.

HOVLAND, Carl, JANIS, Irving, KELLEY, Harold. **Communication and persuasion**. New Haven, CT: Yale University Press. 1953.

KUBISZEWSKI, Ida; NOORDEWIJER, Thomas; COSTANZA, Robert. Perceived credibility of internet encyclopedias. **Computers & Education**, v. 56, n. 3, p. 659-667, 2011. Disponible en: <http://www.uvm.edu/gjee/pubpdfs/Kubiszewski_2011_Computers_and_Education.pdf>. Acceso en: ene. 2012.

LIM, Sook. College students' credibility judgments and heuristics concerning Wikipedia. **Information Processing and Management**, v. 49, n. 2, p. 405-419, 2013. Disponible en: <<http://www.sciencedirect.com/science/article/pii/S0306457312001239>>. Acceso en: mar. 2013.

LIU, Ziming. Perceptions of credibility of scholarly information on the web. **Information Processing and Management**, v. 40, n. 6, p. 1027-1038, 2004. Disponible en: <<http://www.sciencedirect.com/science/article/pii/S0306457303000645>>. Acceso en: mar. 2009.

LIU, Ziming; HUANG, Xiaobin. Evaluating the credibility of scholarly information on the web: a cross cultural study. **The International Information & Library Review**, v. 37, n. 2, p. 99-106, 2005. Disponible en: <<http://www.sciencedirect.com/science/article/pii/S1057231705000299>>. Acceso en: mar. 2009.

LUCASSEN, Teum; SCHRAAGEN, Jan Maarten. The influence of source cues and topic familiarity on credibility evaluation. **Computers in Human Behavior**, v. 29, n. 4, p. 1387-1392, 2013. Disponible en: <<http://www.sciencedirect.com/science/article/pii/S0747563213000393>>. Acceso en: ago. 2013.

METZGER, Miriam; FLANAGIN, Andrew; ZWARUN, Lara. College student web use, perceptions of information credibility, and verification behavior. **Computers & Education**, v. 41, n. 3, p. 271-290, 2003. Disponible en: <<http://www.comm.ucsb.edu/faculty/flanagin/CV/MetzgerFlanaginandZwarun2003%28C&E%29.pdf>>. Acceso en: mar. 2009.

OCDE. Organización para la Cooperación y el Desarrollo Económico. **PISA 2009 Results: Students On Line: Digital Technologies and Performance**, v. 6, 2011. Disponible en: <http://www.oecd-ilibrary.org/education/pisa-2009-results-students-on-line_9789264112995-en> Acceso en: mar. 2011.

PERELMAN, Flora et al. Construcción de criterios de selección en internet en situaciones didácticas: un estudio con estudiantes de escuela primaria. **Revista Interamericana de Psicología/Interamerican Journal of Psychology**, v. 43, n. 3, p. 496-506, 2009. Disponible en: <<http://www.redalyc.org/articulo.oa?id=28412903009>>. Acceso en: ene. 2010.

RIEH, Soo Young; BELKIN, Nicholas. Interaction on the web: scholars' judgment of information quality and cognitive authority. In: ANNUAL MEETING ASIS - AMERICAN SOCIETY FOR INFORMATION SCIENCE AND TECHNOLOGY, 63., Chicago, 12-16 nov. 2000. **Proceedings of the...** Silver Spring, MD: American Society for Information Science, 2000. p. 25-33.

RIEH, Soo Young; BELKIN, Nicholas. Understanding judgment of information quality and cognitive authority in the www. In ANNUAL MEETING ASIS - AMERICAN SOCIETY FOR INFORMATION SCIENCE AND TECHNOLOGY, 63., Chicago, 12-16 nov. 2000. **Proceedings of the...** Silver Spring, MD: American Society for Information Science, 2000. p. 279-289.

TERRA, Ana Lúcia; SÁ, Salvina. Tudo o que vem à rede é peixe? A credibilidade da informação na web. In: CONGRESO NACIONAL DE BIBLIOTECÁRIOS, ARQUIVISTAS E DOCUMENTALISTAS, 11., **Actas...** Lisboa, 2012. Disponible en: <<http://www.bad.pt/publicacoes/index.php/congressosbad/article/view/441/0>>. Acceso en: mayo 2013.

TSENG, Shawn; FOGG, B. J. Credibility and computing technology. **Communications of the Association of Computing Machinery**, n. 42, p. 39-44. 1999. Disponible en: <<http://cacm.acm.org/magazines/1999/5/7905-credibility-and-computing-technology/fulltext>>. Acceso en: mar. 2009.

VEGA DÍAZ, Guadalupe; ROJAS-DRUMMOND, Sylvia. ALFIN para alumnos de educación primaria: propuesta de modelo y perspectivas de aplicación. In: HERNÁNDEZ SALAZAR, Patricia (Coord.). **Tendencias de la alfabetización informativa en Iberoamérica**. México: UNAM, Centro Universitario de Investigaciones Bibliotecológicas, 2012 p. 159-208.

WATHEN, Nadine; BURKELL, Jacquelyn. Believe it or not: factors influencing credibility on the web. **Journal of The American Society for Information Science and Technology**, v. 53, n. 2, p.134-144, 2002. Disponible en: <<http://onlinelibrary.wiley.com/doi/10.1002/asi.10016/pdf>>. Acceso en: mar. 2009.

Received on July 14th, 2013

Approved on December 12th, 2013

Marina Kriscautzky holds a Master's in Science and specializes in Education Research in Departamento de Investigaciones Educativas del Centro de Investigación y Estudios Avanzados (Department of Educational Research of the Center for Research and Advanced Studies) in Mexico. Coordinator of Technology for Education of Dirección General de Cómputo y de Tecnologías de Información y Comunicación (General Directorate of Computer and Information and Communication Technology) of Universidad Nacional Autónoma de México.

Emilia Ferreiro is an emeritus researcher of Sistema Nacional de Investigadores de México (Mexico National System of Researchers). Many of her books have been translated in Brazil, from *Psicogênese da Língua escrita* (with Ana Teberosky, Artes Médicas, 1986) to *O ingresso na escrita e nas culturas do escrito* (Cortez, 2013).