Using Internet videos to learn about controversies:

Evaluation and integration of multiple and multimodal documents by primary school students

Salmerón, L.[1], Sampietro, A.[2], & Delgado, P.[1]

[1] ERI Lectura, University of Valencia, Spain

[2] Jaume I University, Spain

Manuscript accepted for publication at Computers & Education, 01/2020

Address for correspondence: Ladislao Salmerón. University of Valencia, Spain. Avd. Blasco Ibáñez, 21. 46010 Valencia, Spain. E-mail: Ladislao.Salmeron@valencia.edu

Abstract

In many Internet videos authors appear in front of the camera to present their particular view on a topic. Given the high consumption rate of Internet videos by teenagers, we explored the pros and cons of using these videos to learn about complex topics, compared to learning from textual web pages. Specifically, we studied how 207 primary school students (grades 4-6) evaluated and integrated multiple and multimodal web pages (text or video) while learning about the pros and cons of bottled water. Results showed no major role of modality in students' source memory, as measured by citations in their responses to an integration question and their memory for sources. Nevertheless, modality exerted a strong influence on students' beliefs about the topic because, after the study period, they defended the views described in the videos more than those presented in texts. Finally, modality tended to influence students' integration, with participants who learned from two textual webpages including almost twice as many inferences in their responses as those who learned from two videos. We discuss the results in light of current theories of evaluation and integration of multimodal information and (shallow) digital reading, and we elaborate on the pros and cons of using Internet videos in Primary School.

Keywords: multimodal information; Internet videos; Shallowing hypothesis; multiple document comprehension; Primary School education

1. Introduction

The rise of the Internet and the expansion of digital devices are rapidly changing the nature of literacy in the 21st century (Bråten, Braasch, & Salmerón, 2018). Students and citizens in general are not only expected to master traditional literacy skills, such as reading printed texts, but also to be efficient in navigating hyperlinked documents (e.g. Web pages, Wikipedia), integrating information from diverse sources (e.g. a section in a textbook and a blog entry) and evaluating the quality of the information (e.g. whether to trust a report in social media) (Kiili et al., 2018b; Leu, Kinzer, Coiro, Castek, & Henry, 2013; Rouet, 2006; Salmerón, Strømsø, Kammerer, Stadtler, & van den Broek, 2018). Literacy research has focused on understanding how students use textual, and to some extent pictorial, information, but there is still a lack of research on how students use, integrate, and evaluate information from videos for educational purposes (Ainsworth, 2018; Andresen, Anmarkrud, & Bråten, 2019; Andresen, Anmarkrud, Salmerón & Bråten, 2019; List, 2018; List & Ballenger, 2019; Smith, Kiili, & Kauppinen, 2016). This situation contrasts with the high consumption rate of Internet videos by teenagers, both for leisure and for informational purposes (Eurostat, 2016).

Internet videos have great potential to support students' learning. Authors from any background can produce videos to present their views on specific issues, from documenting the effects of climate change on their village to discussing the benefits of a vitamin bought on the Internet. Having access to different perspectives is particularly relevant in learning about socio-scientific controversies because they stimulate the construction of an unbiased representation of the issue (Braasch & Bråten, 2017).

This variety also comes with a price because videos, like virtually any content on the Internet, can be created by authors with varying degrees of expertise on the topic discussed (from topic experts to laypersons) and with different motivations (from informing the population to selling a product) (Bråten et al., 2018). Thus, students must evaluate their claims before accepting the information conveyed in the videos as true. In addition, to obtain

a complete understanding from different perspectives, information from videos has to ultimately be integrated with information from other videos or other modes, such as texts or pictures (i.e. multimodal reading) (Andresen et al., 2019). Internet videos not only convey an oral message, but they also visually present the author. Therefore, students' evaluations of the view presented through this format, and its integration, may differ from evaluations of documents that only present textual information.

Therefore, because the risks may exceed the benefits, we urgently need to comprehend how students in primary school evaluate and integrate multiple and multimodal information before we can safely promote such educational practices in schools. Before specifying the rationale and predictions of our study, we discuss relevant theoretical assumptions and prior empirical work.

1.1. Evaluation of multimodal and multiple information

Multiple and multimodal reading refers to the combined use of text, images, and videos to build knowledge (Ainsworth, 2018). For example, this type of reading would take place when a student is learning about climate change in her science class and has to integrate a description of the phenomena explained in text in the textbook with information from a short Internet video in which a scientist from a local institution explains the recent effects of climate change in the region. Learning from multiple documents is conceived of as a multi-layer process where readers ideally comprehend the different views on the issue (i.e. individual situation models), identify and evaluate each position based on credibility cues such as the competence level of the author, and integrate these views into a balanced representation (i.e. intertext model) (Rouet & Britt, 2011). Two main evaluation processes can occur: an automatic evaluation that judges the validity of the claims, regardless of readers' motivation, to support their existing view on the topic; and a strategic evaluation that takes place as long as the reader is induced by the requirements of the task or internally motivated (Richter & Maier, 2017).

To what extent do the two evaluation processes depend on the modality through which the information is presented? Modes, a term initially used within systemic-functional linguistics (Halliday, 1994) and social semiotics (Hodge & Kress, 1998; Kress, 2010), refer to resources for meaning-making (e.g. speech, images, writing, gestures, color, sound, etc.) used by people in different contexts (Kress, 2010). Within a social semiotics perspective, representation and communication draw upon a variety of modes, which can potentially contribute to meaning (Bezemer & Jewitt, 2010). Central to this perspective is the concept of mode affordance, defined as the ability of each mode to easily convey a particular information aspect. Meanings from one mode are always interconnected with the meaning from other modes, which may have pedagogical implications (cf. Jewitt, 2008), especially in the current digital age (Kress, 2003). Internet videos are multimodal artifacts, as they consist of the multimodal semiotic resources of images, speech, gesture, and captions, among others. Watching and listening to the author while he or she gives a message provides a richer pragmatic context than just reading the message because it involves not only the voice, but also the physical context where the communication takes places. This context potentially provides additional credibility cues conveyed by particular modes that may serve as affordances for further evaluation. For example, a receiver could easily identify different characteristics of the source, such as a doctor's white coat, which could be used to reflect on and interpret the author's view.

Recent evidence suggests that automatic validation of information takes place when a learner is presented with information from audiovisual modes, as previously established with textual information (Piest, Isberner, & Richter, 2018). Piest et al. (2018) used a Stroop task to examine epistemic validation of audiovisual information. Participants saw an image and listened to a statement that was congruent, or not, with what was represented visually, and they immediately saw the words "Correct" or "False". Their task was to indicate whether they found the claims to be true or false. Both children and adults tended to be slower when

congruent audiovisual information was followed by the word “False” rather than the word “Correct”, whereas the opposite occurred with incongruent audiovisual information. This effect was evident in the first block of studies, when students had not acquired any strategic knowledge to cope with the Stroop effect. Authors interpreted this pattern as evidence that people automatically evaluate the validity of audiovisual information to the same extent as with textual information.

Regarding strategic evaluation processes, evidence from persuasion studies (Schroeder & Epley, 2015, 2016; Schroeder, Kardas, & Epley, 2017) and developmental psychology (Eyden, Robinson, Einav, & Jaswal, 2013; Einav, Robinson, & Fox, 2013) provides discrepant accounts. Results from persuasion studies suggest that strategic validation of audiovisual information may differ from that of textual information because people use specific cues such as speech or the author’s picture to qualify messages. Schroeder et al. (2017) analyzed to what extent people attribute mental capacities to authors of polarized messages depending on the format in which their messages are displayed: text, voice, or videotaped. The hypothesis was that human speech provides paralinguistic cues that make an author seem mentally capable. Because messages provided in text alone lack these cues, people can more easily dehumanize the author, especially if they disagree with the belief expressed in the message. In one study (Schroeder et al., 2017, Exp. 1), people read, listened to, or watched authors give a series of polarized messages, and immediately after that, they rated the authors’ human uniqueness (using a 6-item scale related to higher-order cognition and intellectual competence) and human nature (using a 6-item scale related to emotional experience and interpersonal warmth). When participants disagreed with the belief expressed in the message, they rated the author of videos and audios as having more human-uniqueness and human-nature traits than the authors of the text. However, when they agreed with the message, participants rated the authors as having human traits to a similar degree, regardless of the format. The humanizing effect of adding the author’s voice to a text or a video has been

replicated in studies where participants rated a job candidate's appeal or evaluated whether a message came from a human or a machine (Schroeder & Epley, 2015, 2016). Further evidence of the benefits of watching authors comes from computer-mediated communication. In the study by Burgoon, Stoner, Bonito and Dunbar (2003), dyads discussed different topics through computers. They either interacted with their partner through videoconference or through text alone. After each discussion, participants rated their partner's trustworthiness. Participants in the audiovisual condition rated their partners as more trustworthy than those in the text condition.

Developmental studies present an alternative view, proposing that young students may give a particular status to text material, beyond what they give to audiovisual information. Investigating the persuasiveness of print, Einav and colleagues suggest that young children, as soon as they can decode written words, may grant a special status to print that might keep them from adopting a critical stance towards print sources (Corriveau, Einav, Robinson, & Harris, 2014; Einav et al., 2013; Einav, Rydland, Grøver, Robinson, & Harris, 2018; Eyden, et al., 2013). Eyden et al. (2013), for example, found that young readers more easily adopted unexpected suggestions presented to them in written form than those that were orally conveyed to them. This selective trust in printed information arises as students are introduced to formal reading in class, and not before (Einav et al., 2018), suggesting that young students may transfer the authority given to the formal school context where they learn to read to the print message itself.

Note that both hypotheses, the 'humanizing voice' and the 'persuasiveness of print', focus particularly on differences between formats, but not between (more or less trustworthy) sources with a similar format. As described above, an important aspect of learning from multiple documents is identifying and evaluating the information sources in order to qualify their claims (Rouet & Britt, 2011). Because audiovisual information makes authors more salient, participants watching a particular view of a controversy on a video and reading

another view in a text may use the more salient source of information to qualify the view discussed to a greater extent than the view described in the text.

1.2. Integration of multimodal and multiple information

In analyzing the effects of multiple representations on learning, we must identify the purpose of these representations because they can be used for several distinct goals, including complementary roles, constraining interpretations, and constructing deeper understanding (Ainsworth, 1999). Internet videos where authors appear in front of the camera, compared to texts, add a salient visual representation of the information source to the message. Therefore, such videos support complementary information because they can exploit the differences between representations to clearly express the source (i.e. author of the video) as a visual entity. Internet videos can potentially be used for other purposes, such as constructing deeper understanding by depicting abstract representations, as is frequently done in math tutorials. However, in our study, we will focus on a particular type of Internet videos, whose main difference from the textual representation is the audiovisual representation of the author.

A major challenge when learning from multiple representations is the process of translating the information conveyed into the same code in order to be able to integrate them (Ainsworth, 1999; Mayer, 2005). The cognitive theory of Multimedia learning (Mayer, 2005) assumes that textual and pictorial information are processed through two non-interfering channels that can be exploited to overcome the limited processing capacity of the working memory. Information from both channels can be integrated as long as students actively use their prior knowledge to combine them. Thus, according to the multimedia principle, people learn better from text and pictorial information than from text alone. From this perspective, Internet videos may allow better integration of source information (i.e. who said what), due to the fact that sources and messages are processed through different working memory channels Note that this prediction is also consistent with the notion of mode affordance, as the visual

presence of authors in the videos may facilitate students' source identification (Bezemer & Jewitt, 2010).

An alternative view comes from media studies that have focused on the particular way people interact with information from Internet sources. The shallowing hypothesis suggests that most of our current interactions with digital media consist of quick episodes driven by immediate rewards (e.g. number of "likes" in response to an uploaded Instagram video) (Annisette & Lafreniere, 2017). Processing of digital information becomes superficial, and understanding complex information found on the Internet becomes a challenge because students need to focus in order to construct a coherent representation of the message displayed. Distractions associated with quick interactions may disrupt this process, resulting in an incomplete and less coherent representation (cf. Delgado, Vargas, Ackerman & Salmerón, 2018). Because most students tend to use streaming videos for entertainment purposes, even when they are asked to use the Internet only for educational purposes (Malamud, Cueto, Cristia, & Beuermann, 2019), we expect Internet videos to also be processed in a shallow way. By contrast, because students use texts for learning at school on a daily basis, we expect this effect to be less pronounced on textual web pages. Thus, according to the shallowing hypothesis, learning from Internet videos will result in lower comprehension and integration than when reading textual pages.

Although most previous studies have used textual documents to study integration processes, there is an emerging interest in understanding the comprehension and integration of multiple videos. Recently, List and colleagues (Lee & List, 2019; List, 2018; List & Ballenger, 2019) compared the effects of format (text or video) on comprehension and integration of non-conflicting documents. The studies followed a similar procedure. Different groups of undergraduate students read or saw two documents about endangered species, presented either in a video or in a text corresponding to a transcript of the narration in the videos, and later responded to comprehension questions for each document (Lee & List, 2019;

List & Ballenger, 2019), and/or wrote an argument about the similarities and differences between species described in each document (List, 2018; Lee & List, 2019). Across the three studies, the results yielded different patterns. Comprehension of single documents was similar across mediums in List and Ballenger (2019), but it was higher in the video condition in Lee & List (2019). Integration across documents was similar across mediums in List (2018), but it favored the texts condition in Lee & List (2019). Although this research is still preliminary, the pattern of results suggests that videos maybe particularly detrimental for more cognitive demanding integration processes (cf. Annisette & Lafreniere, 2017).

Andresen et al. (2019) investigated how adolescents with or without dyslexia integrated different perspectives of the effects of sun exposure. Documents included information presented in text, pictures, and video. Students' responses to an open integration question indicated that students without dyslexia, compared to those with dyslexia, included more information from textual sources.

Both Andresen et al. (2019) and List's studies (Lee & List, 2019; List, 2018; List & Ballenger, 2019) used multiple multimodal documents with similar (high) degrees of credibility. The question remains of how students would integrate conflicting information from multimodal sources with different credibility. In such situations, solving the controversy (i.e. deciding which is true) requires prior background that is not always available to readers (Stadtler & Bromme, 2014), but they can still attempt to integrate the views by relying on the evaluation processes. Typically, adult readers will defer to expert sources to resolve the contradictions within documents (Clark, Wegener, Habashi, & Evans, 2012; Thomme & Bromme, 2016; Tobin & Raymundo, 2009). However, the author's expertise has been found to be less relevant for secondary school students when they are familiar with the topic at hand (Bråten, McCrudden, Stang Lund, Brante, & Strømsø, 2018; McCrudden, Stenseth, Bråten, & Strømsø, 2016). In relation to this, a major challenge when trying to integrate conflicting documents is the fact that readers tend to selectively focus on the document that supports their

existing beliefs (Salmerón, Kammerer, & Delgado, 2018). Even when readers access belief-inconsistent documents, they tend to be less well understood and retained than belief-consistent ones (text-belief consistency effect, e.g., Maier & Richter, 2013). Consequently, readers tend to fail to correctly characterize and integrate the alternative views into their representation of the topic.

This effect may be dependent on the reader's evaluation of the sources because it may be more difficult to ignore a belief-inconsistent message coming from an expert and neutral author (Clark et al., 2012). To the extent that Internet videos may increase the salience of sources, a belief-inconsistent message from an expert and neutral source depicted in a video may be more difficult to ignore than an equivalent text document.

1.3. Primary school students and multiple documents

As children have increased access to multiple sources on the Internet, there is a growing interest in studying primary school students' skills in critically evaluating and integrating information from multiple documents (Macedo-Rouet, Braasch, Britt, & Rouet, 2013; Paul, Cerdán, Rouet, & Stadtler, 2018; Paul, Stadtler, & Bromme, 2019; Salmerón, Macedo-Rouet, Rouet, 2016). Few studies have analyzed primary school students' ability to handle multiple textual documents.

Primary school students are able to identify expert's occupations from familiar topics. For example, fourth grade students correctly answer questions such as "Who probably knows best what is broken in a car?" (Paul et al., 2018, Exp. 1). But they struggle to use such information to evaluate multiple documents. Macedo-Rouet et al. (2013) found that when reading short controversial accounts of familiar topics, fourth and fifth graders correctly identify more knowledgeable sources (e.g. a veterinarian versus a woman who loves dogs). But when they must draw a conclusion from the controversy, they tend accept the view provided by expert and neutral sources to the same extent as views supported by less knowledgeable sources (Paul et al., 2018; Exp. 2). This is particularly true when students

must analyze a complex controversy, such as when the positions describe not mutually exclusive claims (e.g. in the case of the controversy for or against cereal intake: "Medical doctor advises against the intake since the cereals contain sugar" vs. "Owner of supermarket advises for the intake since the cereals contain oats") (Paul et al., 2019).

Such inefficient use of source information may be partially due to the fact that primary school students have a naïve representation of expert sources. Salmerón et al. (2016) reported that fifth and sixth graders tend to accept claims from knowledgeable authors more often when they support their arguments based on their own experience (e.g. "I had this problem in the past") than when they backed their views with academic evidence (e.g. "I read the solution in a medical encyclopedia") (Salmerón et al., 2016, Exp. 2).

In sum, research suggests that although primary school students can identify accounts from expert sources when reading multiple textual documents, they struggle to apply this knowledge to correctly evaluate the views reported in the documents.

1.4. The current study

Building on theoretical assumptions and empirical work discussed in the previous sections, the present study analyzes primary school students' evaluations and integrated comprehension of the controversy "Tap or bottled water", after reading and watching different web pages with different views on the issue. Participants were randomly assigned to four conditions, in which we crossed the modality (text or video) with the point of view (in favor of or against bottled water). The conditions in which the controversy was introduced, via two texts or two videos, allowed us to compare the effects of modality, which is comparable to the design by List (2018). In the other two conditions, a particular view was supported by a specific modality (e.g. text defending bottled water and video defending tap water). Comparing these conditions allowed us to test the potential influence of the modality on shaping students' beliefs about the controversy. To the best of our knowledge, no previous study has used this design.

Regarding source memory and use, we propose that, because Internet videos provide affordances in the visual mode, which allow an easy identification of the author, source memory and source references in the students' responses should be higher than those for equivalent texts (cf. Bezemer, & Jewitt, 2010). This effect should be particularly strong for aspects visually conveyed in the videos, such as the author's profession, to the extent that it can be inferred from the clothing and the setting.

Regarding source trust, on the one hand, the 'Humanizing voice' hypothesis suggests that students would more often adopt views conveyed in videos than those conveyed in text because human voices would make the sources more trustworthy. On the other hand, the 'Persuasiveness of print' hypothesis makes the opposite prediction, with texts having a special trustworthy status in primary school.

Finally, regarding multiple document comprehension, the multimedia effect suggests that a controversy described in videos could enhance integrated understanding to a greater extent than texts. Students may process the source information in videos through a visual channel, which will not interfere with the processing of the message through a linguistic channel, as could occur when both the source and the message are conveyed through text. By contrast, the 'Shallowing hypothesis" predicts that integrated understanding will be better in texts than in Internet videos, due to the fact that students may process the videos more superficially because they are used to consuming them for pleasure rather than for learning.

To isolate variance resulting from our experimental manipulation, we controlled for the potential effects of school, reading comprehension skills, prior knowledge, and interest in the topic because prior studies indicate that they may be linked to the evaluation and integration of multiple documents in early adolescence (e.g., Kiili, Leu, Marttunen, Hautala, & Leppänen, 2018; Macedo-Rouet et al., 2013; Paul et al., 2018).

2. Method

2.1. Participants

Initially, 207 students from 4th, 5th, and 6th grades participated in the study (*M* age = 10.5, *SD* = .99; 52% girls). Students came from three public primary schools from the suburbs of a major city in Spain (in the Spanish system, 6th grade is the last year of primary education). Schools were located in neighborhoods with families from low to medium socio-economic status. At the three schools, literacy was mainly taught using traditional paper materials, and computer instruction focused on learning office applications and access to Internet information. From the original sample, we excluded data from participants with learning disabilities (*n* = 3) and with incomplete data (*n* = 7) from the analyses. The study was approved by each school board, which listed the study as an academic task for the students.

2.2 Materials

2.2.1 Reading comprehension questionnaire

We used the text comprehension subtask of a widely used standardized test of reading comprehension in Spanish (Cuetos, Rodríguez, Ruano, & Arribas, 2007). Students read two expository texts and subsequently answered 10 short open-ended questions about each text. The test includes a combination of literal questions, which require students to recall specific ideas from the text, and inferential questions, which require them to relate information from different parts of the text or from the text and students' prior knowledge. According to Cohen's benchmarks (1988), the test had a questionable reliability in our sample, McDonald's ω = .60.

2.2.2 Frequency of Internet use

Students' use of the Internet was self-reported questionnaire successfully used in prior research (e.g. Delgado et al., 2019). Briefly, this questionnaire asked whether students had Internet connection at home, how long they had been using the Internet, and how many days a week they currently used it. Additionally, they had to indicate what their usual activities were from a set of common Internet activities.

2.2.3 Topic interest measure

Topic interest is conceived as an internal disposition towards the topic and a tendency to engage with the topic (Hidi & Renninger, 2006). To measure participants' topic interest, two items asked about their interest in the topic of Water and their willingness to learn more about it. Participants indicated their level of interest by rating each item on a 4-point Likert-type scale. For example, to the item "Would you like to learn more about the topic of bottle and tap water", students could respond: "A lot, something, not much, nothing". Reliability scores on this measure was acceptable, McDonald's $\omega = .77$.

2.2.4 Topic beliefs questionnaire

A 5-item measure asked about participants' preferences and beliefs about tap and bottled water. Participants reported their beliefs on a 3-point Likert-type scale. For example, to the item "Which is best for health, tap water or bottle water?", participants could respond: "tap water, bottled water, or both are equally good". Reliability scores on this measure was acceptable, McDonalds $\omega = .70$.

2.2.5 Documents

We created three web pages about the topic "Tap or bottled water", displayed in a mock search engine result list. Characteristics of the documents are described in Table 1. Videos and texts only differed in the presence of the author. In both formats, the source of the web page was provided as a logo on top of the web page. In the videos, the author appeared in front of the camera in a single shot. The source's occupation was displayed in a textual label at the beginning of the video. It could also be partially inferred by the scene because the doctor was wearing a white coat, with a bottle of rubbing alcohol on her desk, whereas the sales manager was dressed formally in an office, with a bottle of a popular brand of mineral water on her desk. The videos were designed to minimize their differences to avoid potential confounding effects other than the source characteristics associated to the characters' profession. We tried to keep constant the characters' age, gender, voice tone and speed, body

position and (lack of) gestures. Characters were recorded in a similar setting (an office), and the camera was setup to record them from a similar angle and distance.

-Table 1 and 2 here-

The texts only included a transcript of the speech given in the corresponding videos. The source's occupation was provided at the end of the text, using the same textual label included in the videos. Texts didn't include any picture of the author. To ensure that texts were appropriate for this population, we computed readability scores for each of the three documents using the Flesch-Szigriszt Index (Szigriszt, 1992), which is a version of the classic Flesch Index for texts in Spanish. The mean readability score for the texts was 74.14 (*SD* = 2.30), indicating that the reading material was "easy" according to the INFLESZ scale (Barrio-Cantalejo et al., 2008). This scale distinguishes five levels of text difficulty, ranging from "very difficult" (readability < 40; e.g., undergraduate textbooks) to "very easy" (readability > 80; e.g., primary school textbooks).

Whereas the first introductory web page was identical in all four conditions, the type of document on the web pages describing the controversy varied across conditions (Table 2).

2.2.6 Integration question

After reading or watching the webpages, all the participants responded to an open integration question. Specifically, they were given the following written instruction: "use the information from the webpages you just saw to answer the following question. What is better for health and the environment, drinking bottled water or tap water?" The average length of the responses was 40.6 words (*SD* = 16.9, *min* = 4, *max* = 119). Responses were analyzed in terms of sourcing, comprehension, and position defended.

We first divided each response into idea units. An idea unit contained a main verb that expressed an event, activity, or state (Magliano, Trabasso, & Graesser, 1999). After

segmentation, all responses were coded to indicate whether the ideas contained an explicit reference to source information, including the author's occupation, the institution that hosted the page, and scientific studies referenced in the web pages.

Finally, ideas were coded to identify students' understanding of the topic. Specifically, we distinguished among three types of ideas: single idea paraphrases, intratext inferences, and intertext inferences. Single idea paraphrases included correct statements in which students used their own words to express an idea from one of the documents without changing its original meaning (e.g., "Plastic bottles release substances that are dangerous to health." [from original website or video] / "Bottled water is inside plastic, which gives off substances that are bad for health." [from student essay]). Single-idea paraphrases thus contained accurate and relevant information about tap and bottled water that was presented in the documents. However, because they do not convey a coherent mental representation of the document's content, they represent a superficial understanding of the topic (McNamara & Magliano, 2009; Salmerón, Gil, & Bråten, 2018a).

Segments were coded as intratext inferences if they combined two single-idea paraphrases from one document that were not connected in the document, or if they contained a single-idea paraphrase linked to some information from students' prior knowledge (e.g., "Natural mineral water is bottled immediately after flowing from the spring with great care so that the water does not lose its characteristics." [from original website or video] / "Bottled water is already clean and doesn't carry anything else; it's just water that most people trust." [from student essay]). Because intratext inferences reflect coherent mental representations of documents, they represent an integrated understanding of the topic, at least at the document level (McNamara & Magliano, 2009; van den Broek & Kendeou, 2015; Salmerón et al., 2018a).

Third, segments were coded as intertext inferences if they combined two single-idea paraphrases across two documents (e.g., "Bottled water is better for your health, but the

plastic in bottles is very polluting. Tap water hasn't been filtered, but it doesn't contain plastic. Conclusion: environment/tap water, health/bottled water." [from student essay]). Because intertext inferences reflect coherent mental representations at the level of multiple documents, they represent an integrated, cross-document understanding of the controversy (Rouet & Britt, 2011; Salmerón et al., 2018a). Due to the limited number of intertext inferences, we created a total inference score by adding up the number of intratext and intertext inferences.

We analyzed students' position on the controversy by identifying the view defended in the response, favoring bottled water, favoring tap water, or not favoring a particular view. Participants were identified as defending tap water or bottled water if: a) they explicitly positioned themselves to defend one view (e.g., "I think tap water is better because it doesn't pollute the environment as much as plastic bottles, and plastic bottles can spread bad things."); b) they discussed more positive arguments than negative arguments about that view. Participants were identified as not favoring a particular view if they included a similar number of arguments in favor of and against the two views (e.g., "I don't know, I think it's better both ways, but for health bottled water and for the environment tap water.").

The first and third authors, blind to the conditions, independently scored a random selection of 9% of the sample to test interrater reliability. The coding of students' references to sources yielded a Cohen's Kappa of 1, the understanding of the content .80, and the view defended in the responses 1, thus showing substantial agreement. All disagreements were resolved through discussion between the two raters, and the first author scored the remaining essays according to the same coding systems.

2.2.7 Memory for sources task

We measured students' memory for sources with a recognition task that presented five correct source names from the web pages (the names of two author occupations plus three web providers), along with seven distractors (the names of three author occupations and four

web providers not mentioned in the web pages). We chose a recognition measure to reduce the memory demands of the task, which has been a useful strategy in previous research with students with limited linguistic abilities (Delgado, Avila, Fajardo, & Salmerón, 2019). To calculate each participant's outcome, we used the *A'* discrimination index, a non-parametric statistic considered appropriate to calculate outcomes from *yes/no* tasks. This index computes a measure by taking into account the percentage of hits, the percentage of false-alarms, and response bias (Stanislaw & Todorov, 1999), yielding a minimum value of 0 and a maximum of 1, where .5 represents performance at the level of chance.

2.2.8 Source-to-content links task

We assessed participants' ability to link text information to the corresponding source by means of a source-to-content links task (Kammerer, Meier, & Stahl, 2016; Stang Lund, Bråten, Brante, & Strømsø, 2017; Strømsø, Bråten, & Britt, 2010). Students were given 9 statements that paraphrased ideas included in each of the three web pages (3 statements per page), along with 3 distracting statements. They were asked to link each statement to the corresponding web page or to the tag 'This is not stated on any of the web pages'. For example, students must link the item "Plastic from water bottles releases substances that can be harmful for humans" to the source "Valencian association of consumers". The score was the total number of statements correctly linked to the corresponding source (*max* = 12). McDonald's ω indicated that test reliability was questionable: .62.

2.3 Procedure

The study took place in two sessions lasting 55 minutes each, in groups of 10-15 participants (each individual class was divided in half). In one session, students completed the reading comprehension test and the Internet frequency of use questionnaire in their own classroom. In the second session, which took place in the computer classroom, students worked individually on a desktop computer equipped with headphones. Randomization of conditions across participants was ensured by assigning conditions to individual computers.

Participants first completed the prior knowledge, interest, and prior beliefs questionnaires. Then, they received instruction about how to play and use the Internet videos. At this point, we took special care to ensure that all participants could listen to the videos using their headphones. After solving any technical problems, we introduced participants to the reading task, where they had ten minutes to read and watch the information on the web pages provided to prepare for different comprehension tasks. To ensure a similar level of exposure to the material across conditions, all the students were told to review the documents if they finished before the time limit. After the reading phase, they completed the integration question, the memory for sources task, and the source-to-content links task. There was no time limit to complete these tasks. All the questionnaires were provided on paper, and students used pencils to respond.

3. Results

3.1 Descriptive and correlational analyses

Descriptive statistics for all the individual measures are displayed in Table 3. On average, students' scores on reading comprehension skills were within the corresponding normality range. Students had high interest in the topic. On the pre-test, most of the students had a positive attitude towards bottled water. Finally, most of them watched Internet videos for fun, and approximately half of them for learning purposes as well.

-Table 3 here-

Data from individual measures and task performance was inspected for normality. Data for interest, source citations, source memory, single ideas paraphrased and total inferences were highly skewed, while the rest of continuous scores were normally distributed. We transformed non-normally distributed raw scores to correct for skewness, which resulted in normal distributions (skewness < 1) for all variables except for source citations, probably due to the high percentage of 0 in such variable. Accordingly, we used parametric analyses for normally distributed data, and non-parametric analyses for scores with skewed distributions

(i.e. source citations) and non-continuous values (i.e. prior topic beliefs and post topic beliefs).

Next, a series of comparisons with condition as independent variable were conducted to identify a priori differences between conditions. Conditions did not differ in terms of reading comprehension skills topic interest (both *Fs* < 1), or prior beliefs (χ^2 (3)= 5.11, *p* = .164) (for descriptive data refer to Table 3).

As Table 4 shows, zero-order correlations between individual and task measures tended to be positive, but rather weak. It is worth noting that scores on the reading comprehension skills test positively correlated with three task measures: source memory, source-to-content links scores, and total inferences on the integration question.Prior topic beliefs was not related to any of the task measures. Finally, task measures were somewhat related to each other. Specifically, inferences on the integration question positively correlated with source-to-content links scores, and negatively with single idea paraphrases.

-Table 4 here-

3.2 Effects on source memory and use

To test our prediction about the effects of the condition on sourcing, we computed a set of mixed model analyses of covariance (ANCOVAs) with condition (two texts, video against bottled water, video pro bottled water, two videos) as fixed independent variable, and source memory as dependent variable. As covariates, we included grade level and scores on the reading comprehension skills test, as they correlated with the dependent variable. School was included as random factor to control for its potential effects. For the variable source citations, we used the non-parametric Kruskal Wallis test. See table 5 for descriptive data and for the main effect of condition.

-Table 5 here-

First, results for source citations in the integration response showed no effect of conditionOverall, students seldom cited sources in their summaries, with approximately one in four citing a source in their responses. Similarly, results for source memory revealed no effect of condition or of school, $F < 1$.. Instead, the covariatesreading comprehension skills, $F(1, 191) = 6.06$, $p = .01$, $\mu^2_p = .03$, and grade, $F(1, 191) = 5.72$, $p = .02$, $\mu^2_p = .02$, exerted a positive small effect, according to Cohen's (1988) effect-size benchmarks. Finally, we computed a fine-grained analysis by considering whether participants identified the professions of the two authors of the videos on the memory for sources task. Chi-square revealed that the number of participants who saw the doctor on the video (two videos and video against bottled water conditions) and identified her profession on the memory for source task was higher (59.81%) than the number who read her message in the text (two texts and video supporting bottled water conditions) (44.21%), χ^2 (1) = 4.79, $p = .029$, Cramer's $V = .16$. This is a small effect according to Cohen's (1988) benchmarks. This effect was not observed for the profession of the sales manager, which was indicated by a similar number of participants on the video (28.43%) and in the text (22.10%), χ^2 (1) = 1.04, $p = .308$.

In sum, the analyses of the sourcing measures indicate that there were no major differences between the document types, although videos seemed to support better memory for some of the authors' professions.

3.4 Effects on source trust

To test the two competing predictions about the effects of the modality on source trust, we measured the adoption of particular beliefs about the controversy ("Humanizing voice" vs. "Persuasiveness of print"). Specifically, we computed chi square tests between conditions on prior topic beliefs (measured by a questionnaire before reading the texts and watching the

videos) and on post topic beliefs (measured by the view defended on the integration question). At pretest, topic beliefs were strongly biased towards bottle water, regardless of the condition, χ^2 (3) = 17.93, p = .006, V = .30 (see Figure 1 for descriptive data). Follow up chi square tests revealed significant differences within each condition (two texts: χ^2 (2) = 17.82, p < .001, V = .30; video against bottled water: χ^2 (2) = 34.12, p < .001, V = .42; video supporting bottled water: χ^2 = 37.73, p < .001, V = .44; two videos: χ^2 (2) = 48.73, p < .001, V = .50). In all four conditions, students initially defended the view that bottled water is better than tap water to a high extent, as indicated by big effect sizes according to Cohen's (1988) benchmarks.

At posttest, topic beliefs, as identified in the position defended on the integration question (support bottled water, tap water, or both), changed across conditions, χ^2 (3) = 17.35, p = .008, V = .30 (see Figure 1 for descriptive data). This change, however, was dependent on the condition. In the two text conditions, a similar percentage of participants held beliefs supporting bottled water, tap water, or both, χ^2 (2) = 1.37, p = .50 (Figure 1, top-left panel). In the condition of the video supporting bottled water, a higher percentage of participants favored bottled water in their responses, χ^2 (2) = 6.68, p = .03, V = .19 (Figure 1, bottom-left panel). By contrast, in the other two conditions, the percentage of participants defending tap water was higher than the percentages for the other two beliefs (video against bottled water: χ^2 (2) = 13.61, p = .001, V = .26, –Figure 1, top-right panel; two videos: χ^2 (2) = 6.52, p = .038, V = .18, –Figure 1, bottom-right panel).

In sum, videos were found to exert a strong influence on students' beliefs about the bottled water or tap water controversy, regardless of their prior beliefs. Across conditions, such effects were big, according to Cohen's (1988) benchmarks.

-Figure 1 here-

3.4 Effects on integrated understanding

To test the two competing predictions about the effects of the condition on integrated understanding (i.e. "multimedia effect" vs. "shallowing hypothesis"), we computed three ANCOVAs with condition (two texts, video against bottled water, video pro bottled water, two videos) as independent variable, and scores on the source-to-content links task, number of single idea paraphrases, and total inferences on the integration task as dependent variables. As covariates, for each analysis we included the variables that correlated with the dependent variable. See table 5 for descriptive data and main effects of condition.

Mixed-model ANCOVA for scores on the source-to-content links task revealed no effect of conditionor of school, $F(1, 191) = 1.56$, $p = .33$. . There were significant effects for the covariates reading comprehension skills ($F(1, 191) = 29.70$, $p <= .001$, $\mu^2_p = .14$) and grade ($F(1, 191) = 6.25$, $p = .013$, $\mu^2_p = .03$). Planned contrasts for the two text and two video conditions revealed no significant differences, $F(1, 191) = 1.78$, $p = .18$.

Results for single idea paraphrases showed no omnibus effect of condition, or school, $F(1, 191) = 2.41$, $p = .26$. A positive significant effect was observed for the covariate topic interest, $F(1, 192) = 8.78$, $p = .003$, $\mu^2_p = .05$. Planned contrasts for the two text and two video conditions revealed no significant differences, $F < 1$.

Finally, the ANCOVA for total inferences revealed no effect of condition, school, $F < 1$, or the covariate grade, $F < 1$, and a positive effect of the covariate reading comprehension skills, $F(1, 191) = 9.27$, $p = .003$, $\mu^2_p = .05$. Planned contrasts for the two text and two video conditions revealed a significant difference, $F(1, 191) = 4.27$, $p = .04$, $\mu^2_p = .02$, small according to Cohen's (1988) benchmarks, with participants in the two-text condition including almost twice as many inferences in their responses as those in the two-video condition (Table 5, lower row).

4. Conclusions

The present study is the first investigation designed to comprehend how students in primary school evaluate and integrate multiple and multimodal (i.e. text and Internet videos) information. Although the pattern of results is far from simple, the results make it possible to weigh the potential risks and benefits of this pedagogical activity for learning. Before elaborating on these practical implications, we will discuss the results in light of current theories of evaluation and integration of multimodal information and (shallow) digital reading.

4.1. Source memory and multimodal information

Authors play a prominent role in the videos used in our study, compared to texts, where authors are only "visible" through a textual label. The presence of the authors in the videos is a material and cultural affordance (cf. Bezemer & Jewitt, 2010) that may allow to easily identify the information source. Therefore, we expected students who watched controversial views on videos to perform better on the memory for sources task, remember more information about the sources, and include a higher number of references to the sources in their responses. Contrary to our hypothesis, results on the number of source references in students' integration question did not vary across conditions. The lack of effects may be partially due to the infrequent inclusion of source references in students' responses, as approximately only one in four students included such references. Similarly, Paul et al., (2018, Exp. 2) reported that 66% of their sample of German fourth graders did not cite any reference in their justifications for selecting a particular web page. Results on the memory for sources task provide only weak support for our hypothesis. No differences were found in students' global memory for sources, a measure that included all the source information available in the documents (i.e., the institution hosting the web page and author's occupation) (cf. Delgado et al., 2019). However, because our videos only make the authors' occupation more salient, we also looked at students' responses to these particular items. Results showed that participants remembered the doctor's occupation more often if they watched the video

(two videos and video against bottled water conditions) than if they read the corresponding text (two texts and video supporting bottled water conditions). This effect was not evident for the occupation sales manager. This differential pattern may be partially due to the varying level of saliency of the videos. Whereas in our videos the doctor could easily be identified because she was wearing a typical white coat, the sales manager could be identified as just a qualified "office clerk" (see screen captures of the videos used in Table 2). Remembering her as a sales manager would have required an additional step because the visual information from the video must be linked to the textual label identifying her occupation. Future studies should replicate the effects, given the small effect sizes for the sourcing effects, and further test this hypothesis by varying the degree of saliency of the display of the source information.

4.2. Source trust and multimodal information

The results allowed us to test two competing predictions about the extent to which students may trust sources depending on the modality (i.e. Internet videos or texts). We measured source trust indirectly by identifying the view defended in students' responses to the integration question about the tap and bottled water controversy. Because most of the students supported bottled water before the study, a critical test was the condition in which the video opposed this view while the text supported it (i.e. video against bottled water). Supporting the 'Humanizing voice' hypothesis (Schroeder & Epley, 2015, 2016; Schroeder et al., 2017), results from this condition indicated that students changed their view towards the one expressed in the video by the expert author (i.e. against bottled water). Similarly, in the video supporting bottled water condition, in which the video supported bottled water while the text opposed it, students more often defended the view expressed in the video. According to the 'Humanizing voice' hypothesis, students perceived the sources on the videos as more human and more capable than those in the text, and so they were more persuaded to change their beliefs about the topic. This pattern of results is difficult to explain with the "Persuasiveness of print" hypothesis (Einav et al., 2013, 2018; Eyden et al., 2013) because sources in texts,

compared to those in videos, were not convincing enough to change students' beliefs. Our data qualify this hypothesis and suggest that the authoritativeness of print that arises in early readers may disappear in subsequent years as students read printed materials as well as different kinds of texts on the Internet outside school. Thus, other contexts would be associated with less authority and truth. For example, when students read novels at home, they may realize that not all printed information represents reality, or when they access online texts such as social media, they may experience the same kind of deception with printed information sources that they experience in face-to-face situations.

Results for the conditions that presented the conflict through the same modality (i.e. text or videos) provide information about the students' ability to critically use sources to evaluate information. In a recent study, Paul et al. (2018) found that fourth graders had difficulties using source characteristics, such as the authors' level of competence or benevolence, to judge the views on a controversial topic presented in text. Similarly, our results for the condition that presented the controversy as two texts show that students did not favor the expert over the commercial view. However, when students were presented with the controversy by means of two videos, the number of students defending the expert view in their responses almost doubled the number defending the commercial view. This effect is surprising, given that young students struggle to identify commercial information from the Internet as incredible, at least when it is presented in text (Kiili, Leu, Marttunen, et al, 2018). To explain our results, we must consider that students as young as four years old are able to judge the credibility of people they interact with based on what they have learned about the adults' abilities (for a review, see Harris, 2012). Because the type of videos used in our study resemble face-to-face situations, in so far as the reader can see the source of the information, while learning in this modality students may apply the advanced evaluation skills they learned in social situations to decide who to trust. As we discuss below, this opens up new

pedagogical approaches to training students to critically assess information online during primary school.

4.3. Integration of multimodal information

Content analysis of the students' responses to the integration question about the controversy studied allowed us to test two competing predictions about the effects of multimodality on integrated understanding. Omnibus analyses revealed no differences between conditions in the number of single-idea paraphrases or the total number of inferences in the responses. Topic interest positively predicted the inclusion of single-idea paraphrases, and reading comprehension skills predicted the inclusion of inferences (Kiili, Leu, Marttunen et al., 2018; Paul et al., 2018). Critical to testing the competing predictions was the planned comparison between the conditions that presented the controversy in one modality, text or video. Supporting the shallowing hypothesis (Annisette, & Lafreniere, 2017), our results indicated that participants who watched the controversy in videos included fewer inferencesthan those reading about the controversy in texts. Although the effect size was small according to Cohen's (1988) benchmarks, the number of inferences in the two text condition almost doubled that of the two video condition. No differences were observed in single idea paraphrases. This effect cannot be attributed to differences in time exposure because all the participants studied the controversy for the same amount of time. Because virtually all the students in our sample used Internet videos for leisure purposes, and far less for learning purposes, they may have approached the video learning task in a shallow mode. This approach resulted in no differences in terms of single idea paraphrases, which are easier to process, although there were differences on the more effortful inferential processing. These results speak to the importance of the modality in text comprehension, and they add to the growing body of research suggesting that as students are increasingly exposed to digital media, often for different purposes, they need to learn how to master their use of media for learning (for a recent review see Delgado et al., 2018). This result contrasts with what was

recently found by List (2018), who reported no differences in integration between a multiple text condition and a multiple video condition in a sample of undergraduate students (see also Merkt, Weigand, Heier, & Schwan, 2011), but is similar to the effect reported by Lee and List (2019). The large difference between these studies in terms of the age of the sample (primary school vs. undergraduate students) keeps us from drawing any strong conclusions about this effect. Potentially, because older students have ample opportunities to learn from Internet videos during high school, this experience may have increased their awareness of the need to actively approach learning tasks through this medium.

4.4. Educational implications for learning with multiple and multimodal information

Our study contributes to the recent interest in studying primary school students' skills in evaluating and integrating information from multiple documents (Macedo-Rouet, et al., 2013; Paul, et al., 2018; 2019; Salmerón, et al., 2016). We identified potential benefits and risks of using multimodal information to learn about controversial topics, and so the promotion of these educational practices in schools should depend on the pedagogical goal. On the one hand, presenting multimodal information from sources with different credibility can promote students' evaluation and use of source information because primary school students (fourth to sixth grade) would be able to differentiate expert sources from commercial bias sources presented as Internet videos. As such, introducing controversies through these videos in late primary school can create a bridge between students' skills in evaluating people's credibility in face-to-face interactions and the lack of strategic evaluation they show on textual documents (Paul et al., 2018). On the other hand, because controversies fully presented through Internet videos hamper inferential processing, they should be avoided if the pedagogical goal is to promote the construction of an integrated understanding of the topic. Instead, presenting the controversy through multimodal sources, as we did in the condition where a sales manager argued in favor of bottled water on text and a doctor criticized it on video, might be a good compromise because it allows students to identify and acknowledge

the view expressed by an expert, without disturbing their integrated comprehension. Nevertheless, the trend to increasingly use Internet videos in primary school is unlikely to decrease, and so future research should explore ways to increase students' active engagement with videos.

4.5. Limitations

Of course, our study has certain limitations. A major problem of our study is that some of the measures used, such as the reading comprehension skills questionnaire and the source-to-content links task, showed questionable internal reliability. The effects of these variables, although not central to our study, should be interpreted with caution. In addition, as other studies found (Paul et al., 2018), primary school students seldom refer to sources in their written responses, which limited the usefulness of this measure. Future research should use a set of more reliable measures, including new measures such as prior knowledge, to fully capture the complex process of attending to, identifying, and using sources such as student interviews (Paul, Macedo-Rouet, Rouet, & Stadtler, 2017) and eye-movement recordings (Salmerón, Gil, & Bråten, 2018b).

In addition, the experimental setup was difficult to implement in the school context, requiring us to test large groups (10-15 students) in one room. Although this is a rather standard way of conducting research at schools, the fact that our students had to use headphones, which were prone to technical problems, and that for the sake of randomization in each class, students saw different videos, produced some disturbance in a few groups. Finally, our results are restricted to a particular topic and only refer to students in the last years of primary school. In the same line, we designed our videos to minimize the differences other than source occupation. As a consequence, characters didn't vary in terms of other potentially relevant modes, such as gestures or voice tone. While we should be cautious not to overgeneralize our results to any multimodal learning environment, we hope our study will

stimulate further research on multiple and multimodal processing using other topics, different populations, and modes.

5. References

Ainsworth, S. (1999). The functions of multiple representations. *Computers & Education, 33,* 131-152.

Ainsworth, S. E. (2018). Multi-modal, multi-source reading: A multi-representational reader's perspective. *Learning and Instruction, 57,* 71-75.

Andresen, A., Anmarkrud, Ø., & Bråten, I. (2019). Investigating multiple source use among students with and without dyslexia. *Reading and Writing*, 32, 1149-1174.

Andresen, A., Anmarkrud, Ø., Salmerón, L., & Bråten, I. (2019). Processing and learning from multiple sources: A comparative case study of students with dyslexia working in a multiple source multimedia context. Frontline Learning Research. *Manuscript in press.*

Annisette, L. E., & Lafreniere, K. D. (2017). Social media, texting, and personality: A test of the shallowing hypothesis. *Personality and Individual Differences, 115,* 154-158.

Barrio-Cantalejo, I. M., Simón-Lorda, P., Melguizo, M., Escalona, I., Marijuán, M. I., & Hernando, P. (2008). Validación de la Escala INFLESZ para evaluar la legibilidad de los textos dirigidos a pacientes. *Anales del Sistema Sanitario de Navarra, 31*, 135-152.

Bezemer, J. & Jewitt, C. (2010). Multimodal analysis: Key issues. In L. Litosseliti (Ed.). *Research methods in linguistics* (pp. 180-197). London, UK: Continuum.

Braasch, J. L., & Bråten, I. (2017). The discrepancy-induced source comprehension (D-ISC) model: Basic assumptions and preliminary evidence. *Educational Psychologist, 52,* 167-181.

Bråten, I., Braasch, J., Salmerón, L. (in press). Reading Multiple and Non-Traditional Texts: New Opportunities and New Challenges. In E. B. Moje, P. Afflerbach, P. Enciso, & N. K. Lesaux (Eds.), *Handbook of Reading Research (Vol. V).* New York: Routledge.

Bråten, I., McCrudden, M. T., Stang Lund, E., Brante, E. W., & Strømsø, H. I. (2018). Task-Oriented learning with multiple documents: Effects of topic familiarity, author expertise, and content relevance on document selection, processing, and use. *Reading Research Quarterly*, *53*, 345-365.

Burgoon, J. K., Stoner, G. A., Bonito, J. A., & Dunbar, N. E. (2003, January). Trust and deception in mediated communication. In System Sciences, 2003. Proceedings of the 36th Annual Hawaii International Conference on (pp. 11-pp). IEEE.

Clark, J. K., Wegener, D. T., Habashi, M. M., & Evans, A. T. (2012). Source expertise and persuasion: The effects of perceived opposition or support on message scrutiny. *Personality and Social Psychology Bulletin*, *38*, 90-100.

Cohen J. (1988). Statistical power and analysis for the behavioral sciences. Hisdale, NJ: Lawrence Erlbaum Associates.

Corriveau, K. H., Einav, S., Robinson, E. J., & Harris, P. L. (2014). To the letter: Early readers trust print-based over oral instructions to guide their actions. *British Journal of Developmental Psychology*, *32*, 345–358.

Cuetos, F., Rodríguez, B., Ruano, E., & Arribas, D. (2007). *PROLEC-R. Batería de evaluación de procesos lectores. Revisada.* Madrid: TEA.

Delgado, P., Avila, V., Fajardo, I., & Salmerón, L. (2019). Training young adults with intellectual disability to read critically on the Internet. *Journal of Applied Research in Intellectual Disabilities, 32*, 666-677.

Delgado, P., Vargas, C., Ackerman, R., & Salmerón, L. (2018). Don't throw away your printed books: A meta-analysis on the effects of reading media on comprehension. *Educational Research Review, 25*, 23-38.

Einav, S., Robinson, E. J., & Fox, A. (2013). Take it as read: Origins of trust in knowledge gained from print. *Journal of Experimental Child Psychology*, *114*, 262–274.

Einav, S., Rydland, V., Grøver, V., Robinson, E. J., & Harris, P. L. (2018). Children's trust in print: What is the impact of late exposure to reading instruction? *Infant and Child Development, 27(6)*, e2102.

Eyden, J., Robinson, E. J., Einav, S., & Jaswal, V. K. (2013). The power of print: Children's trust in unexpected printed suggestions. *Journal of Experimental Child Psychology, 116*, 593–608.

Eurostat (2016). Internet access and use statistics - households and individuals. Retrieved from: http://ec.europa.eu/eurostat/statistics-explained/index.php/Internet_access_and_use_statistics_-_households_and_individuals#Further_Eurostat_information

Halliday, M.A.K. (1994). *An introduction to functional grammar*. London, UK: Arnold.

Harris, P. L. (2012). *Trusting what you're told: How children learn from others*. Harvard University Press.

Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist, 41*, 111-127.

Hodge, R., & Kress, G. (1998). *Social semiotics*. Cambridge, UK: Polity Press.

Jewitt, C. (2008). Multimodality and Literacy in School Classrooms. *Review of Research in Education, 32*(1), 241–267.

Kammerer, Y., Meier, N., & Stahl, E. (2016). Fostering secondary-school students' intertext model formation when reading a set of websites: The effectiveness of source prompts. *Computers & Education, 102*, 52–64.

Kiili, C., Leu, D. J., Marttunen, M., Hautala, J., & Leppänen, P. H. (2018). Exploring early adolescents' evaluation of academic and commercial online resources related to health. *Reading and Writing, 31*, 533-557.

Kiili, C., Leu, D. J., Utriainen, J., Coiro, J., Kanniainen, L., Tolvanen, A., Lohvansuu, K. & Leppänen, P. H. (2018). Reading to learn from online Information: Modeling the factor structure. *Journal of Literacy Research, 50,* 304-334.

Kress, G. (2003). *Literacy in the new media age*. London: Routledge

Kress, G. (2010). *Multimodality. A social semiotic approach to contemporary communication.* New York: Routledge.

Leu, D. J., Kinzer, C. K., Coiro, J., Castek, J., & Henry, L.A. (2013). New literacies: A dual-level theory of the changing nature of literacy, instruction, and assessment. In D. E. Alvermann, N. J. Unrau, & R. B. Ruddell (Eds.), *Theoretical models and processes of reading (6th ed., pp. 1150–1181).* Newark, DE: International Reading Association.

Lee, H.Y., & List, A. (2019). Processing of texts and videos: A strategy-focused analysis. Journal of Computer Assisted Learning. *Manuscript in press*.

List, A. (2018). Strategies for comprehending and integrating texts and videos. *Learning and Instruction, 57,* 34-46.

List, A., & Ballenger, E.E. (2019). Comprehension across mediums: the case of text and video. Journal of Computing in Higher Education. *Manuscript in press.*

Macedo-Rouet, M., Braasch, J. L., Britt, M. A., & Rouet, J. F. (2013). Teaching fourth and fifth graders to evaluate information sources during text comprehension. *Cognition and Instruction, 31,* 204-226.

Magliano, J. P., Trabasso, T., & Graesser, A. C. (1999). Strategic processes during comprehension. *Journal of Educational Psychology, 91*, 615-629.

Maier, J., & Richter, T. (2013). Text belief consistency effects in the comprehension of multiple texts with conflicting information. *Cognition and Instruction, 31,* 151-175.

Malamud, O., Cueto, S., Cristia, J., & Beuermann, D. W. (2019). Do children benefit from internet access? Experimental evidence from a developing country. Journal of Development Economics. *Manuscript in press.*

Mayer, R.E. (2005). Cognitive theory of multimedia learning. In R.E. Mayer (Ed.) *The Cambridge handbook of multimedia learning* (pp. 31-48). New York: Cambridge University Press.

McCrudden, M. T., Stenseth, T., Bråten, I., & Strømsø, H. I. (2016). The effects of topic familiarity, author expertise, and content relevance on Norwegian students' document selection: A mixed methods study. *Journal of Educational Psychology, 108,* 147-162.

McNamara, D. S., & Magliano, J. P. (2009). Toward a comprehensive model of comprehension. *Psychology of Learning and Motivation, 51*, 297-384.

Merkt, M., Weigand, S., Heier, A., & Schwan, S. (2011). Learning with videos vs. learning with print: The role of interactive features. *Learning and Instruction, 21*, 687-704.

Paul, J., Cerdán, R., Rouet, J. F., & Stadtler, M. (2018). Exploring fourth graders' sourcing skills/Un análisis de la capacidad de escrutinio sobre las fuentes de información de los estudiantes de cuarto grado. *Infancia y Aprendizaje, 41*, 536-580.

Paul, J., Macedo-Rouet, M., Rouet, J. F., & Stadtler, M. (2017). Why attend to source information when reading online? The perspective of ninth grade students from two different countries. *Computers & Education, 113*, 339-354.

Paul, J., Stadtler, M., & Bromme, R. (2019). Effects of a sourcing prompt and conflicts in reading materials on elementary students' use of source information. *Discourse Processes, 56,* 155-169.

Piest, B. A., Isberner, M. B., & Richter, T. (2018). Don't believe everything you hear: Routine validation of audiovisual information in children and adults. *Memory & Cognition, 46,* 849–863.

Richter, T., & Maier, J. (2017). Comprehension of multiple documents with conflicting information: A two-step model of validation. *Educational Psychologist, 52,* 148-166.

Rouet, J. F. (2006). *The skills of document use: From text comprehension to Web-based learning.* Mahwah, NJ: Erlbaum.

Rouet, J. F., & Britt, M. A. (2011). Relevance processes in multiple document comprehension. In M. T. McCrudden, J. P. Magliano, & G. Schraw (Eds.), *Text relevance and learning from text* (pp. 19-52). Greenwich, CT: Information Age.

Salmerón, L., Gil, L., & Bråten, I. (2018a). Effects of reading real versus print-out versions of multiple documents on students' sourcing and integrated understanding. *Contemporary Educational Psychology, 52,* 25-35.

Salmerón, L., Gil, L., & Bråten, I. (2018b). Using eye-tracking to assess sourcing during multiple document reading: A critical analysis. *Frontline Learning Research, 6,* 104-122.

Salmerón, L., Kammerer, Y., & Delgado, P. (2018). Non-academic multiple source use on the Internet. In J.L.G. Braasch, I. Bråten, & M. T. McCrudden (eds.), *Handbook of Multiple Source Use (pp. 285-302).* New York: Routledge.

Salmerón, L., Macedo-Rouet, M., & Rouet, J-F. (2016). Multiple viewpoints increase students' attention to source features in social question and answer forum messages. *Journal of the Association for Information Science and Technology. 67,* 2404–2419.

Salmerón, L., Strømsø, H. I., Kammerer, Y., Stadtler, M., & van den Broek, P. (2018). Comprehension processes in digital Reading. In M. Barzillai, J. Thomson, S. Schroeder, and P. van den Broek (Eds.) *Learning to read in a digital world (pp. 91-120).* Amsterdam: John Benjamins.

Schroeder, J., & Epley, N. (2015). The sound of intellect: Speech reveals a thoughtful mind, increasing a job candidate's appeal. *Psychological Science, 26,* 877–891.

Schroeder, J., & Epley, N. (2016). Mistaking minds and machines: How speech affects dehumanization and anthropomorphism. *Journal of Experimental Psychology: General, 145,* 1427–1437.

Schroeder, J., Kardas, M., & Epley, N. (2017). The humanizing voice: Speech reveals, and text conceals, a more thoughtful mind in the midst of disagreement. *Psychological Science, 28,* 1745-1762.

Smith, B. E., Kiili, C., & Kauppinen, M. (2016). Transmediating argumentation: Students composing across written essays and digital videos in higher education. *Computers & Education, 102,* 138-151.

Stadtler, M., & Bromme, R. (2014). The content–source integration model: A taxonomic description of how readers comprehend conflicting scientific information. In D.N. Rapp & J.L.G Braasch (Eds.), *Processing inaccurate information: Theoretical and applied perspectives from cognitive science and the educational sciences (pp. 379 - 402).* Cambridge, MA: MIT Press.

Stang Lund, E., Bråten, I., Brante, E. W., & Strømsø, H. I. (2017). Memory for textual conflicts predicts sourcing when adolescents read multiple expository texts. *Reading Psychology, 38*, 417-437.

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers*, *31*, 137-149.

Strømsø, H. I., Bråten, I., & Britt, M. A. (2010). Reading multiple texts about climate change: The relationship between memory for sources and text comprehension. *Learning and Instruction, 18,* 513-527.

Szigriszt, F. (1992). *Sistemas predictivos de legilibilidad del mensaje escrito: Fórmula de perspicuidad*. Unpublished doctoral dissertation, Universidad Complutense de Madrid.

Thomm, E., & Bromme, R. (2016). How source information shapes lay interpretations of science conflicts: Interplay between sourcing, conflict explanation, source evaluation, and claim evaluation. *Reading and Writing, 29,* 1629-1652.

Tobin, S. J., & Raymundo, M. M. (2009). Persuasion by causal arguments: The motivating role of perceived causal expertise. *Social Cognition, 27,* 105-127.

van den Broek, P., & Kendeou, P. (2015). Building coherence in web-based and other non-traditional reading environments: Cognitive opportunities and challenges. In R. J. Spiro, M. DeSchryver, M. S. Hagerman, P. M. Morsink, & P. Thompson (Eds.), *Reading at a crossroads? Disjunctures and continuities in current conceptions and practices* (pp. 104-114). New York: Routledge.

Table 1

Characteristics of the documents used in the study

Order	Web page	Author	Modality	Main point	Screenshot and URL
Link 1	Educational webpage (wiki)	n.a.	Text + video	Water is important for life. Water needs to be purified before consumption.	 https://mmedia.uv.es/buildhtml/48230
Link 2	Consumer association	Doctor	Text (97 words) or video (25”)	Bottled water can be bad for health. A study from the Minister shows a link between plastic and cancer.	 https://mmedia.uv.es/buildhtml/48229
Link 3	Fontvella (main brand of bottled water)	Sales manager of Fontvella	Text (78 words) or video (25”)	Bottled water doesn’t lose any characteristics. It is good for a healthy diet, and for babies. It is healthier than tap water.	 https://mmedia.uv.es/buildhtml/48231

Table 2

Document type by condition

Page	Two texts	Video against bottled water	Video defending bottled water	Two videos
Educational webpage	Text + video	Text + video	Text + video	Text + video
Consumer Association	Text	Video	Text	Video
Main brand of bottled water	Text	Text	Video	Video

Table 3

Descriptive statistics for measured variables

Variable	*4th grade*	*5th grade*	*6th grade*
N	79	78	40
Reading comprehension skills +	12.5 (2.2)	12.5 (2.1)	13.0 (2.1)
Topic interest +	3.6 (.5)	3.3 (.8)	3.5 (.6)
Prior topic beliefs (in favor of bottled water) ++	79%	90.7%	68%
Watch Internet videos for fun ++	96.3%	92.9%	90.2%
Watch Internet videos to learn ++	45.1%	58.8%	41.5%

Note. + Mean and standard deviation (in brackets). ++ Percentage of students in favor of bottled water.

Table 4

Zero-order correlations between condition and measured variables

Variable	1	2	3	4	5	6	7	8	9	10
1. Grade	-									
2. Reading comprehension skills	.03	-								
3. Topic interest	-.14*	.12	-							
4. Prior topic beliefs +	-.04	-.16*	.00	-						
5. Source citations	-.05	-.05	-.03	.02	-					
6. Source memory	.18*	.18*	.00	.14	.00	-				
7. Source to content links	.18*	.39**	.10	-.04	.00	.23**	-			
8. Single idea paraphrases	-.11	-.03	.22**	-.06	.11	.13	.08	-		
9. Total inferences	.00	.30**	-.01	-.11	.06	.08	.39**	-.41**	-	
10. Post topic beliefs +	.06	-.18*	-.16*	.19**	-.10	.02	-.20**	-.11	-.19*	-

Note. Pearson correlations, except those that involve grade, prior and post topic beliefs, for which we used Spearman rank correlations. + Positive scores mean in favor of bottle water. *$p < .05$, **$p < .01$.

Table 5

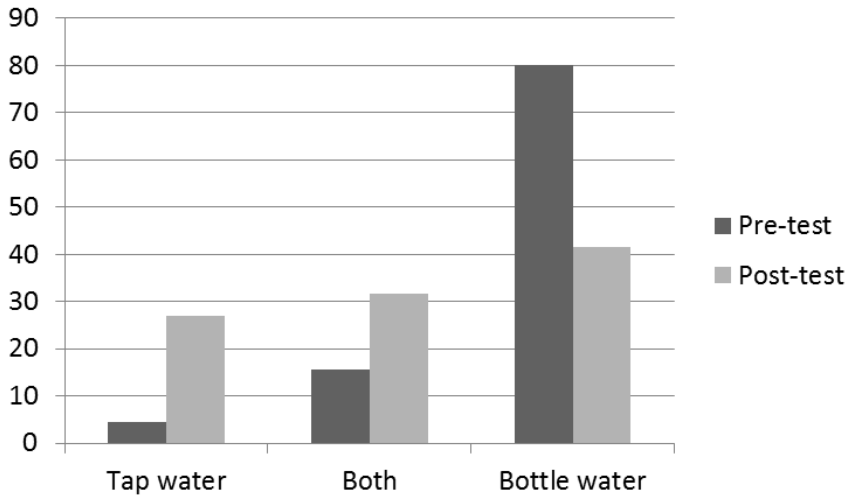
Descriptive statistics (unadjusted means and standard deviations) for measured variables

Variable	Two texts	Video against bottled water	Video defending bottled water	Two videos	Effect of condition
N	45	50	50	52	
Source citations	.23 (.53)	.31 (.62)	.22 (.59)	.10 (.31)	$H(3) = 3.23$, $p = .36$
Source memory	.80 (.17)	.82 (.13)	.82 (.10)	.81 (.14)	$F < 1$
Source to content links	6.18 (2.47)	5.88 (2.83)	6.16 (2.18)	5.81 (2.61)	$F(3, 191) = 1.92$, $p = .13$
Single idea paraphrases	.93 (1.10)	1.13 (1.17)	1.35 (1.45)	1.33 (1.13)	$F(3, 191) = 1.20$, $p = .31$
Total inferences	.58 (.73)	.48 (.68)	.39 (.83)	.31 (.58)	$F(3, 191) = 1.86$, $p = .14$

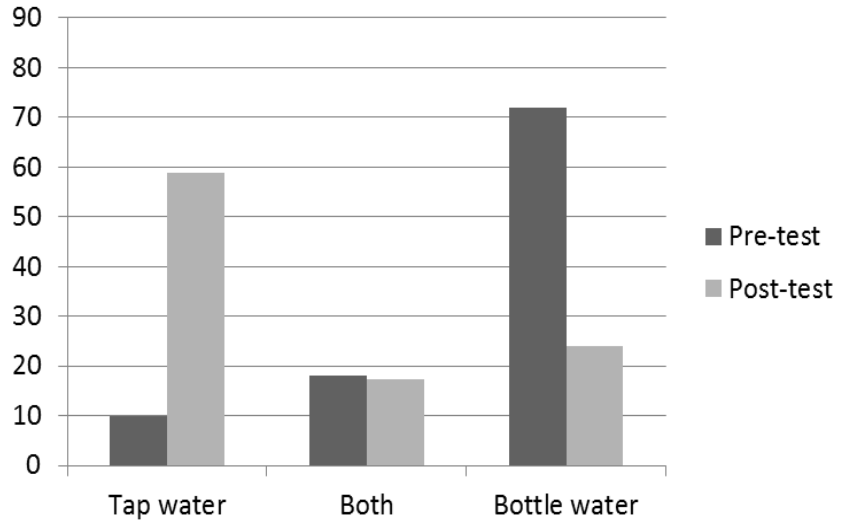
Figure Captions

Figure 1. Percentage of students with particular prior topic beliefs (measured by a questionnaire before reading the texts and watching the videos) and post-task topic beliefs (as measured by the view defended on the integration question) on the controversy, across conditions.

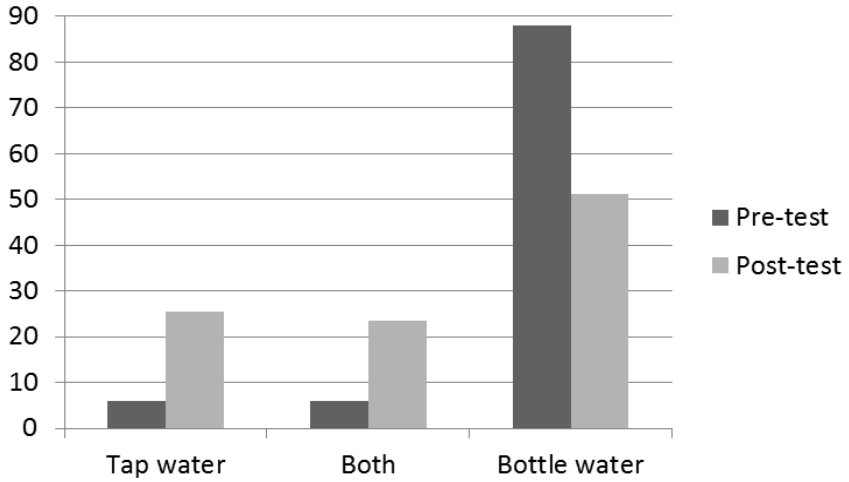
Two texts



Video against bottled water



Video defending bottled water



Two videos