Do you trust this speaker? The impact of prompting on middle-school students' consideration of source when watching conflicting videos

Colin Lescarret (✉ lescarret.colin@gmail.com)

  CLLE: Cognition Langues Langages Ergonomie   https://orcid.org/0000-0002-9227-8127

Julien Magnier

  CLLE: Cognition Langues Langages Ergonomie

Valérie Le Floch

  CLLE: Cognition Langues Langages Ergonomie

Jean-Christophe Sakdavong

  CLLE: Cognition Langues Langages Ergonomie

Jean-Michel Boucheix

  University of Burgundy: Universite de Bourgogne

Franck Amadieu

  CLLE: Cognition Langues Langages Ergonomie

Abstract

In this study, we investigated the impact of prompting on young students' source consideration when watching videos with conflicting information. 262 French 7th graders were confronted to a series of videos in which two speakers (varying in credibility) took opposite stances on the topic of organic farming. Students were either confronted with no prompts (control group), an indirect form of prompting (watching an instructional video on the benefits of sourcing before processing the material), a direct form of prompting (filling out source credibility rating scales during the processing of the material) or a combination of both. While the impact of the instructional video on students' source consideration proved marginal, students who had to fill the source credibility rating scales during the processing of the material better remembered the identity of the speakers (notably in delayed posttest), were more inclined to consider the expert interviewee as the most convincing and to mention interviewees' expertise to justify their judgement. The implications of these results are discussed.

Introduction

Advances in information and communication technologies, and in particular the creation of social networks, have made it possible to disseminate on a large scale unverified or even deliberately falsified information in order to manipulate public opinion (Lazer et al., 2018; Pennycook & Rand, 2021). To find reliable information on a given topic when browsing, it remains therefore necessary to evaluate the quality of the source of information, based on certain criteria such as the source's expertise on the topic, or its intentions (Macedo-Rouet et al., 2019).

A large number of studies have investigated the extent to which young students (elementary and middle-school students) do consider the source of information when reading texts or searching for information online (Braasch et al., 2009; Coiro et al., 2015; Eastin, Yang & Nathanson, 2006; Macedo-Rouet et al., 2013, 2019; Paul et al., 2017, 2018, 2019; Potocki et al., 2020). Overall, these studies indicate that although young students seem to be able to correctly evaluate the quality of a source as long as they process level-appropriate material (Macedo-Rouet et al., 2013; Paul et al., 2018; Potocki et al., 2020), they are less likely to do so as soon as the task becomes more complex (e.g., searching for information online) and the source of information is not salient (Braasch et al., 2009; Coiro et al., 2015; Eastin, Yang & Nathanson, 2006; Macedo-Rouet et al., 2019; Paul et al., 2017). However, prompting students to evaluate the characteristics of the source through, for example, the provision of a criteria grid (Macedo-Rouet al., 2019) appears to greatly improve students' awareness of the source during the processing of the material (Brante & Strømsø, 2018; Kammerer et al., 2016; Macedo-Rouet et al., 2019; Paul et al, 2019; Stadtler et al., 2015).

A limitation of these studies is the exclusive use of textual documents as study material. As noted by Salmerón and colleagues (2020), how young students consider the source when watching videos remain understudied, while video viewing (on streaming) platforms or social networks) is the primary reason for Internet use among the under-20s in France, including children under 12 (IPSOS, 2017). In this study, we

investigated how young students evaluate the source of information when watching videos by exposing seventh graders to videos presenting conflicting information on the topic of organic farming. Two types of sourcing prompts (before viewing, during viewing) were included in the protocol to test their effect on students' source consideration (source memorization, source evaluation). Before presenting the study in more detail, the relevant literature is reviewed.

Do Young Students Have The Skills To Evaluate The Source of Information?

Being able to identify the source of information and to evaluate its credibility are part of the literacy skills required to have a critical look at the information found online (Anmarkrud et al., 2021; Macedo-Rouet et al., 2019; Potocki et al., 2020). Identifying the "source" of information most often refers to locating the author (person or organization) of the document, i.e. the "primary" source, or figuring out who says what within the document, i.e. "embedded sources" (Strømsø et al., 2013; Strømsø, 2017). Evaluating the "credibility" of the source consists in making a judgment on the qualities of the source according to several dimensions (Perloff, 2017; Pornpitakpan, 2004; Stiff & Mongeau, 2016; Wilson & Sherell, 1993), such as source's expertise ("is the source competent enough to talk about this topic?") or source's intentions ("can I trust the source to tell me the truth on the topic?").

Several studies suggest that children are, by the end of elementary school, able to identify the source of information and correctly assess its credibility based on the information provided on its identity – as long as the document is adapted to their level and the task explicitly asks them to consider the source (Macedo-Rouet et al., 2013; Paul et al., 2018; Potocki et al., 2020). Macedo-Rouet and colleagues (2013) presented their participants (French fourth- and fifth-graders) with short texts that featured two characters (one expert on the topic, the other not) providing conflicting statements on several societal topics (e.g., does global warming affect bird migration?). They observed that most participants were able to correctly identify who said what within the texts, and to evaluate which character had the most knowledge about the topic, regardless of their grade level (although fourth-graders had more difficulty justifying their evaluation). Paul and colleagues (2018) also found that their sample of Spanish and German fourth-graders were able to correctly identify the source of information and evaluate its credibility (expertise, trustworthiness) when exposed to short stories adapted to their level.

While these studies highlight young students' adequate sourcing skills when they are confronted with simple tasks (processing of age-appropriate short texts), students' ability to consider the source of information and assess its credibility in more complex tasks (e.g., multiple-documents processing, online information-seeking) is much more queried by the literature (Braasch et al., 2009; Coiro et al., 2015; Eastin, Yang & Nathanson, 2006; Macedo-Rouet et al., 2019; Paul et al., 2017). After exposing elementary school students (3rd to 5th grade) to several webpages varying in authors' expertise and graphic design, Eastin and colleagues (2006) found that the graphic design of the webpages (static vs. dynamic) had more impact on students' evaluation of the authors' expertise than the textual presentation of the authors' qualifications. Coiro and colleagues (2015) observed that 7th graders, while capable to identify the author of a webpage during their search, based their judgments on authors' expertise on vague or superficial

criteria (e.g., that the page provides a graphic). Macedo-Rouet and colleagues (2019) also reported that, when confronted to webpages with quality issues, middle-schoolers (grade 7 to 9) were unlikely to detect issues related to author competence. We can outline from these findings that if young students seem capable to identify the author of certain claims found online, they might encounter more difficulties in evaluating its credibility, and thus tend to assess the quality of the information provided based on superficial cues or criteria.

In response to these difficulties, several studies have tested the influence of instructional activities to improve students' awareness of source when dealing with multiple documents or searching the internet (Brante & Strømsø, 2018). In particular, some studies have shown that the adjunction of incentives to consider the source when completing the task at-hand ("sourcing prompts", Brante & Strømsø, 2018) improves young students' consideration of source information.

Sourcing Prompts As A Mean To Improve Young Students' Consideration Of Source

In an interview study conducted with 9th grade students, Paul and colleagues (2017) found that the lack of incentives to take in account the source was a common justification used by students to explain why they were unlikely to consider the source when reading or searching for information online. According to the RESOLV model (Rouet et al., 2017), this finding suggests that in the absence of explicit instructions to consider source information when performing the task, young students are unlikely to consider the processing of source information as a task demand and thus to integrate it in their task model spontaneously. As a result, several studies have tested the extent to which prompting students to consider source information when completing the task, through a priori reading instructions (Paul et al., 2019; Stadtler et al., 2015) or the completion of a reading grid (Kammerer et al., 2016; Macedo-Rouet et al., 2019), improves young students' source awareness.

Stadtler and colleagues (2015) exposed middle schoolers to a series of online articles regarding the potential health hazards of aspartame consumption. Before reading the articles, half of the participants received a document telling them that not everyone agrees on the issue of aspartame, that some people may have a biased opinion on the issue and therefore to be careful about the source of the articles they were going to read. The results of the study show an effect of prompting on all the measures performed, as participants who were encouraged to consider the source of the articles while reading recalled the sources better and were more inclined to cite the sources in a subsequent argumentative essay. Kammerer and colleagues (2016) asked ninth graders to list on a worksheet the arguments given by four websites varying in quality on the topic of the effect of electromagnetic waves on health. They observed that, when promptings students to pay attention to the quality of the website by dividing the worksheet in categories related to the type of website (scientific, journalistic, personal blogs or commercial), students were more inclined to consider the websites of poor quality as the least credible. Results on younger students are scarcer, but Paul and colleagues (2019) also reported a positive effect of sourcing prompts

on citation of sources in elementary students' argumentative writing (although no effect was found on source recall).

Overall, studies that have examined the effect of sourcing prompts on students' source consideration reports findings in favor of the adjunction of sourcing prompts to improve students' source-awareness, regardless of educational level (Brante & Strømsø, 2018). In line with theories of purposeful reading (e.g., the goal-focusing model of relevance - McCrudden & Schraw, 2007), the hypothesis favored to explain these positive results is that sourcing prompts acts as relevance instructions, i.e. direct students to consider source information as relevant for the task-at-hand (Paul et al., 2019). As such, students are more likely to generate a reading goal oriented towards the localization and processing of source information (Rouet et al., 2017).

A limitation of the studies on sourcing prompts, however, is the focus on textual documents as study material. How young students consider source information when processing videos has been the subject of very little study, despite their ubiquity in students' daily lives (Salmerón et al., 2020). On one hand, processing source information when watching videos might be challenging for young students since this information (e.g., the identity of the speaker) is displayed only for a definite period of time, compared to texts where source information is permanent (Merkt et al., 2011; Merkt & Schwan, 2014). Insofar as students tends to process videos in a passive, linear fashion (Merkt et al., 2011), sourcing prompts may be particularly useful for directing students' attention to information about the source that may otherwise be overlooked by the flow of transient information.

On the other hand, several studies suggest that individuals might be more inclined to consider the source of information when watching videos rather than when reading texts (Chaiken & Eagly, 1983; Booth-Butterfield & Gutowski, 1993, Schroeder et al., 2017), since the combination of sound and image "humanizes" the communicator of a given message and provide additional cues about his/her identity (physical appearance, tone of voice, etc.). After exposing elementary school students to either texts or videos on the topic of bottled water, Salmerón and colleagues (2020) found, for example, that students who were exposed to the videos were (slightly) better at remembering the source of information than students confronted to the texts. Yet, if video is a medium that increases attention to the source compared to text, then the impact of sourcing prompts may be weaker in the context of video viewing than that observed when reading textual documents.

However, these hypotheses remain in the realm of conjecture in the absence of studies that have tested the impact of sourcing prompts in the context of video viewing.

Present Study and Hypotheses

In this study, we investigated whether sourcing prompts would impact young students' consideration of source (considered here as the identity of the people speaking on the screen, i.e. embedded sources) when processing videos with conflicting information. We exposed a sample of seventh grade French

students (average age of 13 years) to fictional video interviews in which two speakers made conflicting statements on the topic of organic agriculture ("Will organic farming be able to feed the entire world population in 2050?"). Source credibility was manipulated by presenting each speaker either as a researcher in field (credible source) or a consumer interrogated while shopping (non-credible source).

Two types of sourcing prompts were included within the procedure to test their impact on students' consideration of source when watching the videos. As a form of "a priori" prompting, half of the students' videos had to watch, before the interviews, an instructional video introducing (1) what a information source was, (2) how to locate it within a webpage or a video, and (3) how to assess its credibility based on two criteria: its expertise on the topic and its intentions (Sourcing Prompt 1: Instructional Video). As a form of "concomitant" prompting, half of the students were asked to assess the credibility of the speakers on screen *while* watching the interviews, using source credibility rating scales embedded in the video player (Sourcing Prompt 2: Embedded Rating Scales).

Because both type of prompts (before and during the processing of the material) have been found to improve source consideration (Brante & Strømsø, 2018), we did not make specific assumptions regarding which type of prompts would be more effective to increase students' source awareness in the context of our experiment. However, we were interested in testing, in an exploratory approach, whether differences would emerge in measures of source consideration based on the type of prompt to which students were exposed. To this aim, the two types of sourcing prompts were manipulated using a fully-crossed between-subjects experimental design, such as one group of students was not exposed to the sourcing prompts (control group), two groups were confronted with one type of prompts but not the other, and a final group was exposed to both types of prompts.

After watching the interviews, students' memorization of the sources was assessed via a recognition task asking students had to correctly recognize the identity of the speakers seen in the videos in a list containing distractors. This test was repeated one week later (delayed posttest). Assuming that the sourcing prompts would increase students' consideration of the identity of the speakers included in the videos, we expected that students confronted with the sourcing prompts would outperform the control group on the recognition task, at both measurement times (Hypothesis 1).

Following the recognition task, students had to rate the credibility of the speakers seen in the videos on two dimensions: perceived expertise, i.e. the perceived level of topic-knowledge or skills of the communicator, and source trustworthiness, i.e. the perceived degree of honesty and goodwill of the communicator (Perloff, 2017). We expected students to rate the speaker presented as a researcher more credible (i.e. more expert and trustworthy) than the speaker presented as a consumer (Hypothesis 2a) and that such difference would be stronger amongst students confronted to the sourcing prompts (Hypothesis 2b). Finally, students had to indicate which speaker they found the most convincing, and justify in writing why. We expected that the students confronted with the sourcing prompts would be more likely to view the credible speaker (i.e., researcher) as the most convincing (Hypothesis 3a) and that they

would more often cite the source's expertise as a justification for their choice (Hypothesis 3b) than the other students.

Method

Sample

Two hundred and sixty-two French seventh graders participated in the study (M_{age}=13 years old, SD = 0.36). Out of this sample, 51.9% of students considered themselves as girls (N = 136). 37.8% of participants studied in middle-schools located in rural area (N = 99), and 62.3% (N=163) in middle-schools located in the center of a large city.

The middle-schools in which the participants were studying enrolled in the study on a voluntary basis, after a call for participation issued to all middle-schools in the Region of Toulouse (South-West of France). Students participated in the study as part of their school curriculum, after obtaining their consent as well as their parents' consent to participate.

Experimental Design

The hypotheses of this study were tested via a 2 (Instructional Video) *2 (Embedded Rating Scales) between-subjects experimental design. The factor "Instructional Video" included the modalities: No Video (N=134), Video Watched in Pretest (N=128). The factor "Embedded Rating Scales" comprised the modalities: Without ERS (N=130), With ERS (N=132). Students were randomly assigned to the experimental groups.

Table 1 indicates the sample size of the four experimental groups resulting from this manipulation.

Table 1 Experimental Groups

Embedded Rating Scales	Instructional Video	
	No Video	Video Watched in Pretest
Without ERS	66	64
With ERS	68	64

Material

Students watched six video extracts of fictitious interviews ($M_{duration}$=49 seconds per video, ET=4.7 sec.) on the topic "Will organic food be able to feed the entire world population in 2050?".

Three of the videos were extracts of an interview with a source arguing that organic agriculture could feed the planet in 2050 and that a transition to 100% organic agricultural production would have positive consequences for the consumer and the environment (*favorable* source); the other three of an interview with a source arguing that this change in agricultural production was not feasible and would have negative consequences (*unfavorable* source). Each of the interviewees gave their opinion on three topics (with one video per theme and per interviewee, for a total of six extracts): (1) the level of productivity of organic farming; (2) the need to change diet to ensure sufficient yield; and (3) the cost to the consumer of a transition to 100% organic farming. Table 2 provides the arguments given by the interviewees on each theme, according to their position (favorable/unfavorable).

Table 2 Arguments Presented By The Interviewees According To Their Position (Favorable/Unfavorable)

	Unfavorable source	Favorable source
Theme 1 = Yield	*Interview extract 1* = Organic farming does not have a sufficient level of productivity to feed the entire world population and switching to fully organic farming by 2050 would require increasing the number of agricultural plots.	*Interview extract 2* = Organic agriculture is able to achieve similar levels of productivity to conventional agriculture through the use of certain farming techniques, which are more environmentally friendly.
Theme 2 = Need to change diet	*Interview extract 3* = Eating less meat means an increased consumption of vegetable proteins to the detriment of animal proteins, which are also necessary for the human body to function.	*Interview excerpt 4* = One third of the world's cereal production is now used to raise livestock. By reducing our consumption of meat, this share will decrease and switching to 100% organic agriculture by 2050 becomes possible.
Theme 3 = Cost to the consumer	*Interview extract 5* = Organic food is more expensive for the consumer because it requires more labour. If by 2050, organic food is the only alternative, then less well-off consumers will find it difficult to eat properly.	*Interview extract 6* = The share of the household budget allocated to food has been falling sharply since the 1950s. Organic products are more expensive, but of better quality and this price is fairer because it allows farmers to be better paid.

The expertise of the interviewees was manipulated by varying their identity (name/occupation). For half of the students (47.7%, N=125), the favorable source was presented as a researcher in the field, i.e., "*Laurent Degain, researcher in agronomy, AgroParisTech (Paris)*", and the unfavourable source as a consumer interviewed while shopping; i.e., "*Christophe, 34 years old*" (Pair 1). For the other half (52.3%, N=137), the favourable source was presented as a consumer interrogated while shopping and the unfavourable source as a researcher in the field (Pair 2). The video playback time was identical for the two pairs (*Total Time*=288 seconds).

Information about the identity of the interviewees was displayed in a banner presented at the bottom of the screen for 5 seconds at the beginning of each extract. Furthermore, the interviews were shot in an

environment congruent with the identity assigned to the interviewees (office for interviewees identified as researchers, supermarket for interviewees identified as consumers), as represented in Figure 1.

The interviewees were actors recruited by an audiovisual agency, and were Caucasian men between 34 and 44 years of age in order to neutralise potential effects of attractiveness and identification at source (DeBono & Harnish, 1988; Pornpitakpan, 2004; Stiff & Mongeau, 2016).

Sourcing Prompts

Instructional Video

Half of students had to watch, during the pretest phase, an instructional video on how and why evaluate the source of information when browsing. The video lasted 6 minutes and 21 seconds and was created with the help of the teachers taking part in the research project, using the PowToon© video creation software.

The video presented, in a first part, what is a source of information and how to identify it when reading an online article or watching a video on the internet (e.g. on YouTube). In a second part, the video insisted on the need to assess the credibility of the source of information and two criteria were provided for this evaluation: the source's expertise on the topic and the source's intentions. Each criterion was defined and guidelines were given for their evaluation. Appendix A further describes the content of the video and provides extracts of the script.

After watching the video, students had to answer a 7-items multiple-choice questionnaire covering the major topics addressed in the video (Ex: *Once you have identified the source of an item of information, what question should you ask yourself to determine whether the information you have found is credible? (1) Can I trust this source of information? (2) Is this information source famous? (3) Does this information source write or speak well? (4) I don't know.* If students answered one of the questions incorrectly, an alert indicated that their answer to that question was wrong and they could not validate the quiz, to ensure that all students had integrated the main points of the video after watching it.

Embedded Rating Scales

In order to encourage students to pay attention to the credibility of the interviewees during viewing, half of students had to answer source credibility ratings scales *during* the viewing of the videos.

After 10 seconds of viewing each excerpt, a question appeared on the screen above the player, aiming to direct students' attention to the credibility of the interviewees speaking on the screen. Clicking on the question paused the video and displayed the response modalities (in the form of a bipolar scale, *see* Figure 2). If students had not answered the question before the video ended, the answer modalities

were automatically displayed and students were asked to give their answer to move on to the next video. The question was displayed after 10 seconds so as not to interfere with the processing of the identity of the source (banner displayed for 5 seconds at the beginning of each extract).

The questions targeted: (1) perceived expertise ("*How expert do you think this person is on the topic?*", bipolar scale scored from -3 "Not at all expert" to +3 "Very much expert") and (2) perceived trustworthiness ("*How much do you trust the information this person gives you?*", bipolar scale scored from -3 "I do not trust it at all" to +3 "I completely trust it"). The question targeting perceived expertise was displayed during the first extract of each source, the question targeting perceived trustworthiness during the second.

Measures

Prior Knowledge

Students' level of prior knowledge on the topic, was assessed using an 8-item multiple-choice questionnaire, with one point for each correct answer (total score from 0 to 8). E.g.: "Nowadays, *how much of the world's cereal production is allocated to feeding livestock (cows, pigs, sheep, etc.)?* (a) *Half of the production.* (b) *One third of production.* (c) *10% of production.* (d) *Don't know.*" The internal consistency of the questionnaire evaluated by Cronbach's alpha was acceptable (α=.75).

Source Recognition

Students' recall of the sources seen in the videos was tested via a forced-choice recognition task. The screenshots of the two interviewees seen in the videos were displayed as recognition cues (without the indication of their name and profession) with the mention "Person A", "Person B". Students had to recognize the correct identity provided in the video (name + occupation) of each interviewee from a list containing the correct answer and five distractors (one of which gave the correct occupation of the source, but not the name). Answers for each source were scored 0 if both the name and profession were falsely recognized, 1 if the correct profession was recognized but not the name and 2 if both the name and profession were correctly recognized. Scores for each source were then summed to provide a total recognition score ranging from 0 to 4.

Source Credibility

After completing the recognition tasks, students were asked to assess the credibility of each interviewee seen in the videos using a selection of items from McCroskey's Authoritativeness and Character Scale (Stiff and Mongeau, 2016). This task was included to the procedure for all participants, even if they had to answer the credibility ratings items during the viewing of the videos. The screenshots of the two

interviewees seen during the videos were provided again as a cue to rate the interviewees' credibility (with their name and occupation replaced by person A and person B).

Perceived Expertise

Four items from McCroskey's Authoritativeness Scale were used to measure how far students perceived the interviewees as expert in the topic. For each item, students were asked to indicate their level of agreement with statements such as "*This person seemed to me to be competent enough to deal with this topic*", on a 7-point Likert scale ranging from 1 ("*Strongly disagree*") to 7 ("*Strongly agree*"). The internal consistency of the questionnaire as measured by Cronbach's alpha was excellent (α=0.88).

Perceived Trustworthiness

Four items from McCroskey's Character Scale were used to measure how far students perceived the interviewees as trustworthy. For each item, students were asked to indicate their level of agreement with statements such as *"I trust this person to tell the truth about the topic*", on a 7-point Likert scale ranging from 1 ("*Strongly disagree*") to 7 ("*Strongly agree*"). The internal consistency of the questionnaire as measured by Cronbach's alpha was very good (α=0.85).

Source convincingness

Lastly, students were asked to indicate which interviewee they found the most convincing, with three possible answers: Person A; Person B; Both were equally convincing. Students were prompted to justify their answers in writing ("*Why did you find this person more convincing than the other one? Justify your answer using the text field below*", if the answer *Person A* or *Person B* was selected, and "*Why did you find both persons equally convincing? Justify your answer using the text field below*", if the answer "*Both were equally convincing* " was selected).

Students' justifications were coded by counting the number of explicit mentions of the interviewees' expertise in their answer (e.g., "*he seemed to know a great deal about it*"). Two independent judges blinded to the study hypotheses coded around 25% of data (*N*=65) justifications). The percentage of agreement between the two judges was very high (96% of agreement), and one judge coded the rest of the data.

Procedure

Phase 1 (pre-test). In this phase, students individually completed an online questionnaire under the supervision of their teachers (who were asked not to intervene except to solve a comprehension problem), during school hours. The instructions stressed that the responses were anonymous and that only the researchers conducting the study would have access to their answers. Students were also informed that

the questionnaire was neither a test nor an evaluation, and were therefore encouraged to answer the questions as honestly as possible.

The questionnaire included (1) the collection of demographic data (first name, age, gender, middle school) (2) the MCQ questionnaire aimed at assessing their level of prior knowledge on the topic.

Half of the students had, after the end of the questionnaire, to watch the instructional video on how to identify and evaluate the source of information (*see above*, section Sourcing Prompts) followed by the MCQ questionnaire aimed at verifying their understanding of the video.

Phase 2 (Viewing Phase / Immediate Posttest). At the beginning of the viewing phase, students were told that they were going to see a series of videos on the topic "Will organic farming be able to feed the entire world population in 2050?", and that they would have to give their opinion on the topic after viewing. Students watched the videos individually with headphones, in their classroom, in the presence of their teacher and an experimenter present to give instructions.

Students viewed the videos using a video player designed for the experiment, which recorded the playback time of each video, in seconds. The player displayed a play/pause button, a progress bar, a volume bar, and a button to toggle the videos to full screen. Bookmarks to access each video were displayed on the left side of the screen (*see* Figure 3). Students had 20 minutes to watch each video at their own pace and, if necessary, watch them again.

Students who had to answer the credibility ratings scale embedded in the video player (*see* Sourcing Prompts) were given the additional instruction that a question would appear on the screen during the viewing of each of the videos and that they could answer it at any time by clicking on the question. If the student had not answered the question at the end of the video, the answer modalities were automatically displayed and students had to give their answer to move on to the next video.

After viewing the videos, all students were asked to complete a second online questionnaire that included: (1) the recognition task aimed at assessing students' memorization of the sources (*see* Measures), (2) the perceived credibility ratings scale (expertise/trustworthiness) of the interviewees. Finally, students were asked to indicate which interviewee they found most convincing and to justify their answer in writing.

Phase 3 (Delayed Posttest). One week later, students were asked to complete a final online questionnaire including the same recognition task than in Phase 2 in order to consider potential long-term changes in the memorization of the sources.

Results

Appendix B summarises the means and standard deviations observed on each measure of the study. Unless otherwise specified, examination of the kurtosis and skewness coefficients of the distributions supported the use of parametric tests to analyse the results of the study, with all coefficients remaining

within acceptable limits to validate the hypothesis of normality of the distributions (Gravetter & Wallnau, 2014).

Analyses were carried out with Jamovi software version 2.2 (The jamovi project, 2021), using the R packages "afex: Analysis of Factorial Experiments" (Singmann, 2018), "emmeans: Estimated Marginal Means" (Lenth, 2020) and "car: Companion to Applied Regression" (Fox & Weisberg, 2020).

No significant difference between the experimental groups was observed in relation to the age of the participants, $F(3,258)=1.08$, $p=.357$, gender, $\chi^2(3)=0.721$, $p=.39$, or level of prior knowledge, $F(3,258)=1.47$, $p=.224$.

We ran a 2 (Instructional Video: No Video, Video Watched in Pretest) *2 (Embedded Rating Scales: Without ERS, With ERS) *2 (Pair of Sources: Pair 1, Pair 2) between-subjects ANOVA on students' total video playback time to test whether students who were confronted with the sourcing prompts watched the videos for a longer time than the other students. However, there was no main effect of the factor Instructional Video, $F(1,254)=1.716$, $p=.191$, $\eta^2_p=.007$, nor of the factor Embedded Rating Scales, $F(1,254)=0.450$, $p=.503$, $\eta^2_p=.002$, nor of the factor Pair of Sources, $F(1,254)=0.892$, $p=.346$, $\eta^2_p=.004$.

An Instructional Video*Embedded Rating Scales interaction effect was noted, $F(1,258)=5.872$, $p=.016$, $\eta^2_p=.023$, as students who were confronted only with the ERS watched the videos for a longer time than students that were confronted with both the Instructional Video and the ERS, [456.98:498.04] (ERS only) vs. 95% CI [416.26:458.26] (Instructional Video + ERS), $t(253)=2.673$, $p_{bonferroni}=.048$, $M_{difference}=400.50$, $SE=149.83$. Post-hoc tests did not reveal any other differences between the experimental groups, and no other interaction effect between the factors included in the ANOVA reached significance (all $p>.10$)

Source Recognition

In order to test the impact of the sourcing prompts on students' recognition of the sources seen in the videos (Hypothesis 1), we ran a mixed ANOVA including one within-subjects factor (Experimental Phase: Immediate Posttest, Delayed Posttest) and three between-subjects factors (Instructional Video: No Video, Video Watched in Pretest; Embedded Rating Scales: Without ERS, With ERS; Pair of Sources: Pair 1, Pair 2) on source identity recognition scores. Video playback time was included in the analysis as a covariate factor, to control its potential impact on recognition scores. The assumptions of homogeneity of variances was met (all Levene's test $p >.05$).

Between-subjects effects revealed a main effect of the factor Embedded Rating Scales, $F(1,253)=6.682$, $p=.010$, $\eta^2_p=.026$. Post-hoc tests indicated that the group of students confronted with the ERS outperformed the group of students who were not, 95% CI [2.60:3.65] (Without ERS) vs. 95% CI [2.86:3.91] (With ERS), $t(253)=-2.64$, $p_{bonferroni}=.003$, $M_{difference}=-0.27$, $SE=.10$. These results partially support Hypothesis 1. There was no effect of the factor Instructional Video, $F(1,253)=0.499$, $p=.542$, $\eta^2_p=.001$, nor

of the factor Pair of Source, $F(1,253)=0.372$, $p=.542$, $\eta^2_p=.001$. No interaction effect between these factors reached significance (all $p > .10$).

Within-subject effects indicated a main effect of the factor Experimental Phase, $F(1,253)=4.477$, $p=.035$, $\eta^2_p=.017$. Post-hoc matched comparisons revealed a significant deterioration of recognition scores between immediate and delayed posttest, 95% CI [3.27:3.47] (Immediate Posttest) vs. 95% CI [3.02:3.25] (Delayed Posttest), $t(253)=4.47$, $p_{bonferroni}<.001$, $M_{difference}=0.235$, $SE=.052$. The Experimental Phase*Instructional Video interaction effect did not reach the significance threshold, $F(1,253)=0.374$, $p=.541$, $\eta^2_p=.001$, but the interaction effect Experimental Phase*Embedded Rating Scales proved significant, $F(1,253)=8.994$, $p=.003$, $\eta^2_p=.034$. As shown in Figure 4, post-hoc matched comparisons indicated that while a significant deterioration of recognition score between immediate and delayed posttest was observed in the group of students who were not confronted with the ERS, 95% CI [3.17:3.47] (Immediate Posttest) vs. 95% CI [2.76:3.10] (Delayed Posttest), $t(253)=5.214$, $p_{bonferroni}<.001$, $M_{difference}=0.392$, $SE=.075$, no significant deterioration in scores was found in the group confronted with the ERS, 95% CI [3.28:3.57] (Immediate Posttest) vs. 95% CI [3.18:3.51] (Delayed Posttest), $t(253)=1.049$, $p_{bonferroni}=1.000$, $M_{difference}=0.077$, $SE=.121$.

An Experimental Phase*Instructional Video*Embedded Rating Scales interaction effect was also observed, $F(1,253)=8.008$, $p=.005$, $\eta^2_p=.031$. As shown in Figure 5, students who were confronted with both the ERS and the Instructional Video obtained significantly higher recognition score in the immediate posttest than students only confronted with the ERS ($p=.025$) and students only confronted with the Instructional Video ($p=.007$). However, a tendential deterioration in performance was observed one week later ($p=.051$) - such that their scores did not differ significantly from the group of students only confronted with the ERS in the delayed posttest ($p=.641$). Yet, students who were confronted with the ERS significantly outperformed the other students in the delayed posttest (all $p<.05$),

No other interaction effect between the factors included in the ANOVA reached significance (all $p>.10$).

Embedded Rating Scales

Mixed ANOVAs including Source as a within-subjects factor (2 levels: Researcher, Consumer) and Pair of Sources as a between-subjects factor (2 levels: Pair 1, Pair 2) were run on expertise and trustworthiness ratings to test whether students who had to answer the credibility ratings scale during the viewing of the videos ($N=132$) found the interviewee presented a researcher more credible on the topic than the interviewee presented as consumer. A strong main effect of Source was found on expertise ratings, $F(1,130)=142.89$, $p<.001$, $\eta^2_p=.524$, and on trustworthiness ratings, $F(1,130)=91.47$, $p<.001$, $\eta^2_p=.413$, as students judged the interviewee presented as a researcher to be more expert than the interviewee presented as a consumer, 95% CI [5.06:5.31] (Researcher) vs. 95% CI [2.94:3.42] (Consumer, $t(130)=12.0$, $p_{bonferroni}<.001$, $M_{difference}=2.13$, $SE=.179$, and more trustworthy, 95% CI [5.14:5.61] (Researcher) vs. 95% CI [3.24:3.80] (Consumer),, $t(130)=9.56$, $p_{bonferroni}<.001$, $M_{difference}=1.86$, $SE=.194$.

There was no main effect of Pair of Sources, $F(1,130)=1.54$, $p=.217$, $\eta^2_p=.012$ (expertise), $F(1,130)=1.51$, $p=.222$, $\eta^2_p=.01$ (trustworthiness), nor any Source * Pair of Sources interaction effect, $F(1,130)=3.37$, $p=.069$, $\eta^2_p=.025$ (expertise), $F(1,130)=2.72$, $p=.102$, $\eta^2_p=.020$ (trustworthiness),

Post-Viewing

To examine whether students rated the credibility of interviewees differently based on the sourcing prompts they were exposed to (Hypotheses 2a and 2b), mixed ANOVAs including one within-subjects factor (Source: Researcher, Consumer) and three between-subjects factors (Instructional Video: No Video, Video Watched in Pretest; Embedded Rating Scales: Without ERS, With ERS; Pair of Sources: Pair 1, Pair 2) were performed on perceived expertise and perceived trustworthiness ratings. The assumption of homogeneity of variances was met (all Levene's test $p >.05$).

Perceived Expertise

Between-subjects effects did not reveal any main effect of the factor Instructional Video, $F(1,254)=1.997$, $p=.159$ $\eta^2_p=.008$, nor of the factor Embedded Rating Scales, $F(1,254)=1.212$, $p=.272$, $\eta^2_p=.005$, nor of the factor Pair of Source, $F(1,254)=0.056$, $p=.812$, $\eta^2_p<.001$. No interaction effect between these factors reached significance (all $p>.10$).

Within-subjects effects revealed a strong simple effect of the factor Source, $F(1,254)=182.049$, $p<.001$, $\eta^2_p=.417$. In line with Hypothesis 2a, post-hoc matched comparisons indicated that students judged the interviewee presented as a researcher more expert than the interviewee presented as a consumer, 95% CI [4.99:5.31] (Researcher) vs. 95% CI [3.14:3.47] (Consumer), $t(254)=13.5$, $p_{bonferron}<.001$, $M_{difference}=1.85$, $SE=.137$. There was no interaction effect between the factor Source and the factor Instructional Video, $F(1,254)=2.144$, $p=.144$, $\eta^2_p=.008$, nor between the factor Source and the factor Embedded Rating Scales, $F(1,254)=0.836$, $p=.362$, $\eta^2_p=.003$, which contradicts Hypothesis 2b.

A Source * Instructional Video * Embedded Prompts was observed, $F(1,254)=5.274$, $p=.022$, $\eta^2_p=.020$. As shown in Figure 6, students who were confronted with both the ERS and the Instructional Video tended to judge the researcher more expert (and the consumer less expert) than the other students. However, no post-hoc matched comparisons between the experimental groups reached significance, either on the expertise ratings of the researcher or the expertise ratings of the consumer (all $p_{bonferroni}>.0$).

No other interaction effect between the factors included in the ANOVA proved significant (all $p >.05$).

Perceived trustworthiness

There was no main effect of the factor Embedded Rating Scales, $F(1,254)=0.295$, $p=.587$, $\eta^2_p=.001$, nor of the factor Pair of Source on perceived trustworthiness ratings, $F(1,254)=0.377$, $p=.540$, $\eta^2_p=.001$. An unexpected main effect of the factor Instructional Video was observed, $F(1,250)=12.144$,

$p<.001$, $\eta^2_p=.046$, as students who had watched the instructional video in pretest found the interviewees to be less trustworthy than the other students, 95% CI [4.30:4.57] (Video Watched in Prestest) vs. 95% CI [4.64:4.91] (No Video), $t(254)=-3.48$, $p_{bonferroni}<.001$, $M_{difference}=0.339$, $SE=.097$. No interaction effect between these factors reached significance (all $p>.05$).

A main effect of the factor Source was noted, $F(1,125) = 57.595$, $p<.001$, $\eta^2_p=.185$. In line with Hypothesis 3a, post-hoc matched comparisons confirmed that students judged the interviewee presented as a researcher to be more trustworthy than the interviewee presented as a consumer, 95% CI [4.89:5.20] (Researcher) vs. 95% CI [4.02:4.31] (Consumer), $t(245)=7.59$, $p_{bonferroni}<.001$, $M_{difference}=0.886$, $SE=.117$.

In discordance with Hypothesis 2b, there was no interaction effect between the factor Source and the factor Instructional Video, $F(1,254)=0.073$, $p=.787$, $\eta^2_p<.001$, nor between the factor Source and the factor Embedded Rating Scales, $F(1,254)=2.199$, $p=.139$, $\eta^2_p=.009$, nor between the three factors, $F(1,254)=2.017$, $p=.157$, $\eta^2_p=.008$. A Source * Pair of Sources interaction effect was noted, $F(1,254)=3.967$, $p=.047$, $\eta^2_p=.015$, as students confronted with the unfavorable researcher and the favorable consumer tended to judge the consumer more trustworthy than students confronted with the favorable researcher and the unfavorable consumer, 95% CI [4.11:4.51] (Consumer – Pair 2) vs. 95% CI [3.80:4.22] (Consumer – Pair 1). However, this difference proved not to be significant, $t(245)=1.990$, $p_{bonferron}=.285$, $M_{difference}=0.292$, $SE=.147$.

No other interaction effect between the factors included in the ANOVA proved significant (all $p>.05$).

Source convincingness

Students' choice

Overall, 47.7% of students ($N=125$) judged the interviewee presented as a researcher to be the most convincing. This percentage proved highly similar whether students were confronted with the favorable researcher (Pair 1: 47.2%, $n=59$) or the unfavorable researcher (Pair 2: 48.1%, $n=66$) and there was indeed no significant association between the pair of sources that students saw in the videos and the proportion of students who rated the researcher as the most convincing source, $\chi^2(1)=0.025$, $p=.875$, Cramer's $V=.010$. There was no association between the proportion of students who rated the researcher as the most convincing interviewee and the factor Instructional Video either, $\chi^2(1)=0.228$, $p=.633$, Cramer's $V=.029$, as the percentage of students who considered the researcher to be the most convincing source was similar whether students had seen the instructional video in pretest (49.2%, $n=63$) or not (46.3%, $n=62$).

Conversely, a significant association between the proportion of students who considered the researcher to be the most convincing interviewee and the factor Embedded Rating Scales was found, $\chi^2(1)=22.1$, $p<.001$, Cramer's $V=.291$. Indeed, the proportion of students who found the researcher to be the most

convincing source was higher among students confronted with the ERS (62.1%, *n*=82) than among students who were not (33.1%, *n*=35).

The implementation of a binomial logistic regression including the factors Instructional Video (two modalities: No Video, Video Watched in Pretest) and Embedded Rating Scales (two modalities: Without ERS, With ERS) as predictors, and as an outcome whether or not students' had found the researcher to be the most convincing interviewee (0: other choice; 1: researcher) confirmed that the factor Embedded Rating Scales was a significant predictor of the outcome, *Estimate*=1.202, *SE*=.259, *Z*=4.641, *p*<.001, *OR*=3.326, while the factor Instructional Video was not, *Estimate*=0.139, *SE*=.259, *Z*=0.536, *p*=.592, *OR*=1.149. The model fit measures proved satisfactory, although the percentage of variance explained remained low, $\chi^2(1)$=22.8, *p*<.001, *AIC*=346, R^2_{CS}=.083.These results provide mixed support for Hypothesis 3a.

Justifications

Overall, 50.8% of students (*N*=133) mentioned interviewees' expertise on the topic to justify which interviewee was the most convincing to them. Because the assumption of linearity was not met for the dependent measure "Number of mentions of interviewees' expertise", a non-parametric Kruskal-Wallis ANOVA was run with Experimental Groups as the between-subject factor (4 levels: Control, Instructional Video only, ERS only, Instructional Video + ERS) to test whether students who were confronted to the sourcing prompts had included more mentions of interviewees' expertise in their justifications compared to the other students (Hypothesis 3b).

A significant effect of the factor Experimental Groups was noted, $\chi^2(3)$=24.30, *p*<.001, ε^2=.093. As represented in Figure 7, DSCF pairwise comparisons indicated that students who were only confronted with the ERS included more mention to interviewees' expertise in their justification than students of the control group (*W*=6.011, *p*<.001) and students who were only confronted with the Instructional Video (*W*=3.911, *p*=.029). Students confronted with both the Instructional Video and the ERS also included more mentions to interviewees' expertise in their justifications than students of the control group (*W*=5.425, *p*<.001), and students only confronted with the Instructional Video – although the latter difference was only tendential (*W*=3.375, *p*=.080). There was no significant difference between the group of students only confronted to the Instructional Video and the control group (*W*=2.238, *p*=.389), nor between the group of students only confronted with the ERS and the group of students confronted with both the Instructional Video and the ERS (*W*=-0.372, *p*=.994). These results partially support Hypothesis 3b.

Discussion

In this study, we were interested in how middle-school students consider the source (identity of the person speaking on the screen) when viewing videos presenting conflicting information about a controversial topic ("Will organic agriculture be able to feed the entire world population in 2050?"). Two types of sourcing prompts were included in the protocol, to test their impact on students' consideration of source

when processing the videos: the viewing of an instructional video on how to identify the source of information and evaluate its credibility *before* processing the material (i.e., a priori sourcing prompts), and the completion of source credibility rating scales *during* the processing of the material (i.e., concomitant sourcing prompts). Based on prior studies (Brante & Strømsø, 2018; Kammerer et al., 2016; Macedo-Rouet et al., 2019; Paul et al., 2019; Stadtler et al., 2015), we expected that the sourcing prompts would improve students' memorization of source information and would encourage them to consider the source presented as an expert (i.e. researcher) as more credible and convincing that the source presented as a layperson (i.e., consumer).

However, the two types of prompting tested in this study did not have an equivalent effect on students' consideration of source. On one hand, students who had to answer the source credibility rating scales during the processing of the videos (i.e., concomitant sourcing prompts) did perform better at the source identity recognition task and were more inclined than the other students to consider the expert source as the most convincing. In addition, they mentioned source expertise more often in their justifications as to which source was more convincing than the other students. Conversely, the impact of watching the instructional video before processing the material (i.e., a priori sourcing prompt) proved to be marginal at best, regardless of the measure considered. Such lack of effect was surprising but may be explained by the objectives of the instructional video and its content.

Indeed, although the instructional video did prompt students to consider the source of information when watching videos, it did not specifically indicate that it was a requirement for the task that would follow. In consultation with the teachers involved in the project, the educational video was created with the goal of encouraging students to consider the source in different contexts (searching for information on the internet, viewing videos on streaming platforms) but did not include any reference to the study material, to allow teachers to reuse the video outside of the experiment. However, this means that students were not given a direct prompt in the instructional video to pay attention to the source in the interviews that would follow. We intentionally did not instruct the students to apply what they had seen in the video to the viewing of the interviews that would follow, to see if they would do it spontaneously.

Yet, the lack of a significant effect of the viewing of the instructional video across our measures suggests that students did not make the connection between the content of the instructional video and the following task (watching video interviews and give their opinion on the topic covered), and thus ultimately processed source information in the same way as the control group. Despite watching a video insisting on the need to take the source into account when consulting documents (including videos), students do not seem to have formed a representation of the subsequent task as requiring particular attention to the source and direct instructions were likely still needed for students to integrate the processing of source information in their task model.

Conversely, having to fill source credibility rating scales during the viewing of the interviews directly required to consider source information when processing the material. Compared to watching the instructional video, the embedded credibility rating scales consisted in a more direct form of sourcing

prompt that "forced" the generation of a viewing goal focused on processing source features - in line with the RESOLV model (Rouet et al., 2017) and the goal-focusing model of relevance (McCrudden & Schraw, 2007), insofar as students were required to answer the credibility ratings to watch the next video. The positive impact of the embedded ratings scales on several measures of source consideration (including source identity memorization) suggests that this form of prompting was thus much more successful in increasing the relevance of source information for the task-at-hand, which is the primary objective of sourcing prompts (Paul et al., 2019). These results are promising, in that they suggest, first, that filling out source credibility ratings scale when processing the material is a valid form of sourcing prompts – in line with previous findings (Macedo-Rouet et al., 2019), and second, that sourcing prompts can help improve young students' consideration of source not only when they process texts, but also when they process other types of documents such as videos - as long as these prompts take the form of an explicit incitation to consider source information during the processing of the material. However, some limitations of the study mean that these conclusions should be treated with caution.

Limitations of the study

First, it should be noted that although several effects of the embedded source credibility ratings scales were found on source consideration, these effects remained moderate. This may be the result of the characteristics of our material. Indeed, the video interviews presented the arguments of two sources that directly contradicted each other. Several studies have shown that this characteristic of the material (conflicting arguments from two different sources or more) can itself foster students' consideration of source, by drawing students' attention to who says what within the document (Braasch & Bråten, 2017; De Pereyra et al., 2014; Salmerón et al., 2016). In this, the effects of sourcing prompts investigated in this study may have been reduced by the experimental material used. This limitation invites replication of the study with different material (e.g., videos including convergent, or only partially divergent information) in order to assess the consistency of the effect sizes found.

Another limitation is that this study focused solely on embedded sources, i.e., the identity of the person speaking on screen (Strømsø, 2017). As such, the conclusions of this study related to the effectiveness of sourcing prompts remain limited to the processing of source information within the video itself. However, the "source" of a video may also refer to the person/organization that produced the video or other characteristics of the video (distribution platform, publication date, etc.). No study to date has been able to investigate how young students take these characteristics into account when selecting which video to watch on social media or streaming platforms like YouTube, and the impact of sourcing prompts on young students' selection of videos to watch when browsing constitutes, in that regard, an interesting area for future research.

Finally, the study lacks some online measures to get more information about how sourcing prompts influenced students' attention to source information (i.e., banner presenting speakers' identities at the beginning of videos). Here, we can infer from the recognition test results that the embedded ratings scales led students to pay more attention to the information about source identity. However, this remains

a postulate in the absence of more direct measures of students' processing such as eyetracking. Conducting post-test interviews with students would be another interesting addition to the protocol in order to gain more information about how students perceived (in terms of, for example, usefulness) the two types of sourcing prompts being tested.

Conclusion

Notwithstanding these limitations, this study is amongst the first to investigate how young students process source information when watching videos, and to our knowledge the first to have tested the impact of several types of sourcing prompts on this processing. Although no significant effect of an indirect form of sourcing prompts (viewing of an instructional video encouraging students to pay attention to the source in pretest) was found on students' source consideration, the filling of source credibility rating scales while watching the material proved successful as improving several measures of source consideration (e.g., which source was found the most convincing). As a result, this form of sourcing prompts seems promising to train young students to pay more attention to the source when watching videos online, although the protocol of the study now needs to be replicated with students of different age groups and on varying video material to evaluate the consistency of its effects.

Understanding how young students' process source information when they watch videos and how to improve this processing appears crucial in a technological context where video has become as preponderant as text (if not more so) as an informational medium, and where video-sharing platforms such as social media allow anyone to post videos on any topic they want - including videos with erroneous or manipulated information.

Declarations

Conflicts of interest: The authors declare no conflicts of interest.

Funding: This study was conducted as part of the research project "DIMEDD" which was funded by the French Ministry of Education, Higher Education and Research through the PIA (Future Investment Program) call for projects.

Data availability: The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

References

1. Anmarkrud, Ø., Bråten, I., Florit, E., & Mason, L. (2021). The role of individual differences in sourcing: a systematic review. *Educational Psychology Review*, 1-44.
2. Booth-Butterfield, S., & Gutowski, C. (1993). Message modality and source credibility can interact to affect argument processing. *Communication Quarterly, 41*(1), 77-89.

3. Braasch, J. L., & Bråten, I. (2017). The discrepancy-induced source comprehension (D-ISC) model: Basic assumptions and preliminary evidence. *Educational Psychologist, 52*(3), 167-181.

4. Braasch, J. L., Lawless, K. A., Goldman, S. R., Manning, F. H., Gomez, K. W., & MacLeod, S. M. (2009). Evaluating search results: An empirical analysis of middle school students' use of source attributes to select useful sources. *Journal of Educational Computing Research, 41*(1), 63-82.

5. Brante, E. W., & Strømsø, H. I. (2018). Sourcing in text comprehension: A review of interventions targeting sourcing skills. *Educational Psychology Review, 30*(3), 773-799.

6. Chaiken, S., & Eagly, A. H. (1983). *Communication modality as a determinant of persuasion: The role of communicator salience. Journal of Personality and Social Psychology, 45(*2), 241.

7. Coiro, J., Coscarelli, C., Maykel, C., & Forzani, E. (2015). Investigating criteria that seventh graders use to evaluate the quality of online information. *Journal of Adolescent & Adult Literacy, 59*(3), 287-297.

8. De Pereyra, G., Belkadi, S., Marbach, L., & Rouet, J. F. (2014, August). *Do teenage readers' use source information when faced with discrepant information?* Paper presented at the annual meeting of the Society for Text and Discourse, Chicago, USA.

9. DeBono, K. G., & Harnish, R. J. (1988). Source expertise, source attractiveness, and the processing of persuasive information: A functional approach. *Journal of Personality and Social Psychology, 55*(4), 541.

10. Eastin, M. S., Yang, M. S., & Nathanson, A. I. (2006). Children of the net: An empirical exploration into the evaluation of Internet content. *Journal of Broadcasting & Electronic Media, 50*(2), 211-230.

11. Fox, J., & Weisberg, S. (2020). car: Companion to Applied Regression. [R package]. Retrieved from https://cran.r-project.org/package=car.

12. Gravetter, F., & Wallnau, L. (2014*). Essentials of statistics for the behavioral sciences (8th ed.)*. Wadsworth.

13. IPSOS. (2017, March 14). *Junior Connect' 2017 : les jeunes ont toujours une vie derrière les écrans !* IPSOS.Com. Retrieved April 12, 2022, from https://www.ipsos.com/fr-fr/junior-connect-2017-les-jeunes-ont-toujours-une-vie-derriere-les-ecrans

14. 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.

15. Lazer, D. M., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., ... & Schudson, M. (2018). The science of fake news. *Science, 359*(6380), 1094-1096.

16. Lenth, R. (2020). emmeans: Estimated Marginal Means, aka Least-Squares Means. [R package]. Retrieved from https://cran.r-project.org/package=emmeans.

17. 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*(2), 204-226.

18. Macedo-Rouet, M., Potocki, A., Scharrer, L., Ros, C., Stadtler, M., Salmerón, L., & Rouet, J. F. (2019). How good is this page? Benefits and limits of prompting on adolescents' evaluation of web information quality. *Reading Research Quarterly, 54*(3), 299-321.

19. McCrudden, M. T., & Schraw, G. (2007). Relevance and goal-focusing in text processing. *Educational Psychology Review, 19*(2), 113-139.

20. Merkt, M., & Schwan, S. (2014). How does interactivity in videos affect task performance?. *Computers in Human Behavior, 31*, 172-181.

21. 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*(6), 687-704.

22. Muller, A., Schader, C., Scialabba, N. E. H., Brüggemann, J., Isensee, A., Erb, K. H., Smith, P., Klocke, P., Leiber, F., Stolze, M. & Niggli, U. (2017). Strategies for feeding the world more sustainably with organic agriculture. *Nature Communications, 8*(1), Article 1290.

23. 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. *Journal for the Study of Education & Development , 41*(3), 536-580.

24. 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.

25. 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*(2), 155-169.

26. Pennycook, G., & Rand, D. G. (2021). The psychology of fake news. *Trends in Cognitive Sciences, 25*(5), 388-402.

27. Perloff, R. M. (2017). *The dynamics of persuasion: communication and attitudes in the twenty-first century*. Routledge.

28. Pornpitakpan, C. (2004). The persuasiveness of source credibility: A critical review of five decades' evidence. *Journal of Applied Social Psychology, 34*(2), 243-281.

29. Potocki, A., de Pereyra, G., Ros, C., Macedo-Rouet, M., Stadtler, M., Salmerón, L., & Rouet, J. F. (2020). The development of source evaluation skills during adolescence: exploring different levels of source processing and their relationships (El desarrollo de las habilidades de evaluación de las fuentes durante la adolescencia: una exploración de los distintos niveles de procesamiento de las fuentes y sus relaciones). *Journal for the Study of Education and Development, 43*(1), 19-59

30. Rouet, J. F., Britt, M. A., & Durik, A. M. (2017). RESOLV: Readers' representation of reading contexts and tasks. *Educational Psychologist, 52*(3), 200-215.

31. 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*(10), 2404-2419.

32. Salmerón, L., Sampietro, A., & Delgado, P. (2020). Using Internet videos to learn about controversies: Evaluation and integration of multiple and multimodal documents by primary school students.

Computers & Education, 148, 103796.

33. 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.

34. Singmann, H. (2018). afex: Analysis of Factorial Experiments. [R package]. Retrieved from https://cran.r-project.org/package=afex.

35. Stiff, J. B., & Mongeau, P. A. (2016). *Persuasive communication*. Guilford.

36. Strømsø, H. I. (2017). Multiple models of multiple-text comprehension: a commentary. *Educational Psychologist, 52*(3), 216-224.

37. Strømsø, H. I., Bråten, I., Britt, M. A., & Ferguson, L. E. (2013). Spontaneous sourcing among students reading multiple documents. *Cognition and Instruction, 31*(2), 176-203.

38. The jamovi project (2021). jamovi. (Version 2.2) [Computer Software]. Retrieved from https://www.jamovi.org.

39. Wilson, E. J., & Sherrell, D. L. (1993). Source effects in communication and persuasion research: A meta-analysis of effect size. *Journal of the Academy of Marketing Science, 21*(2), 101-112.

Figures

Figure 1

Banner presenting the identity of the interviewees

Figure 2

Embedded Rating Scales
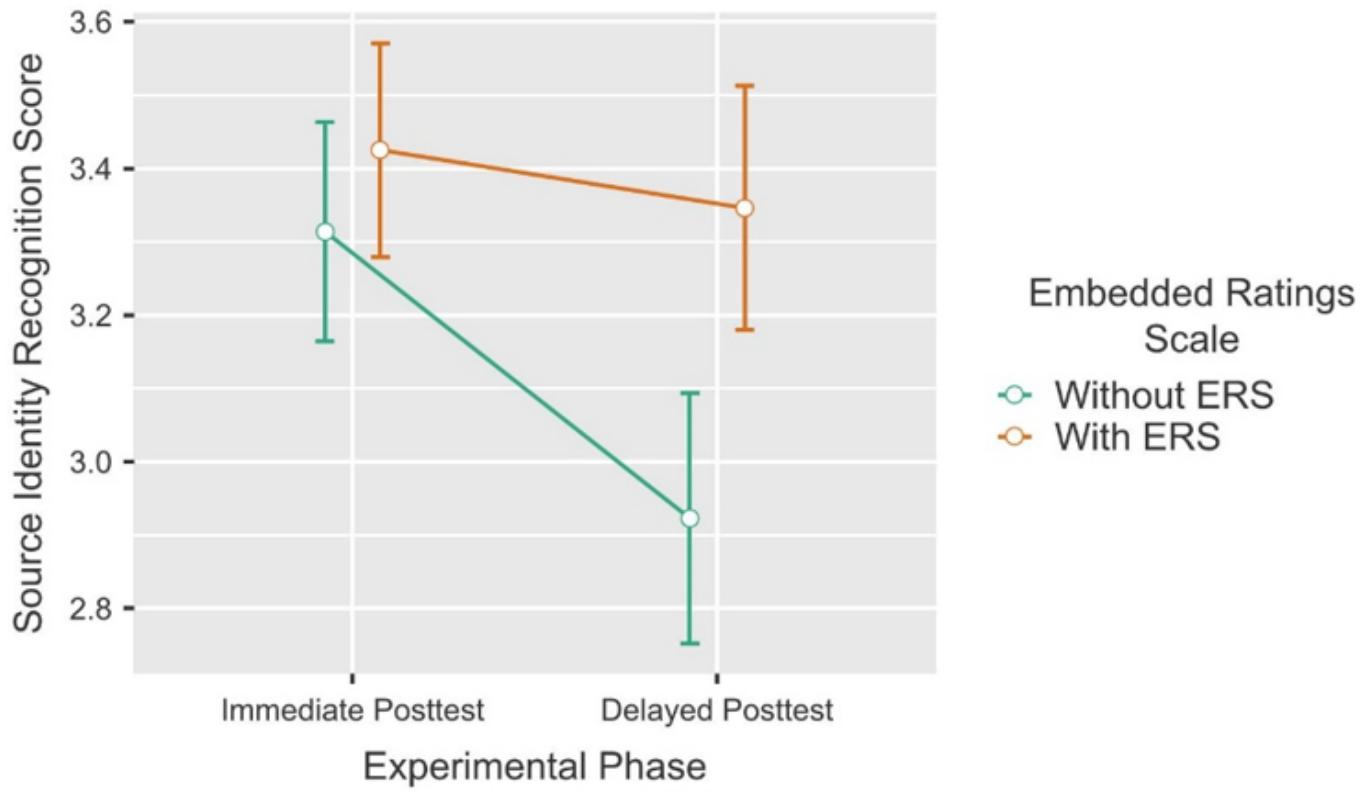

Figure 3

Screenshot of the viewing page

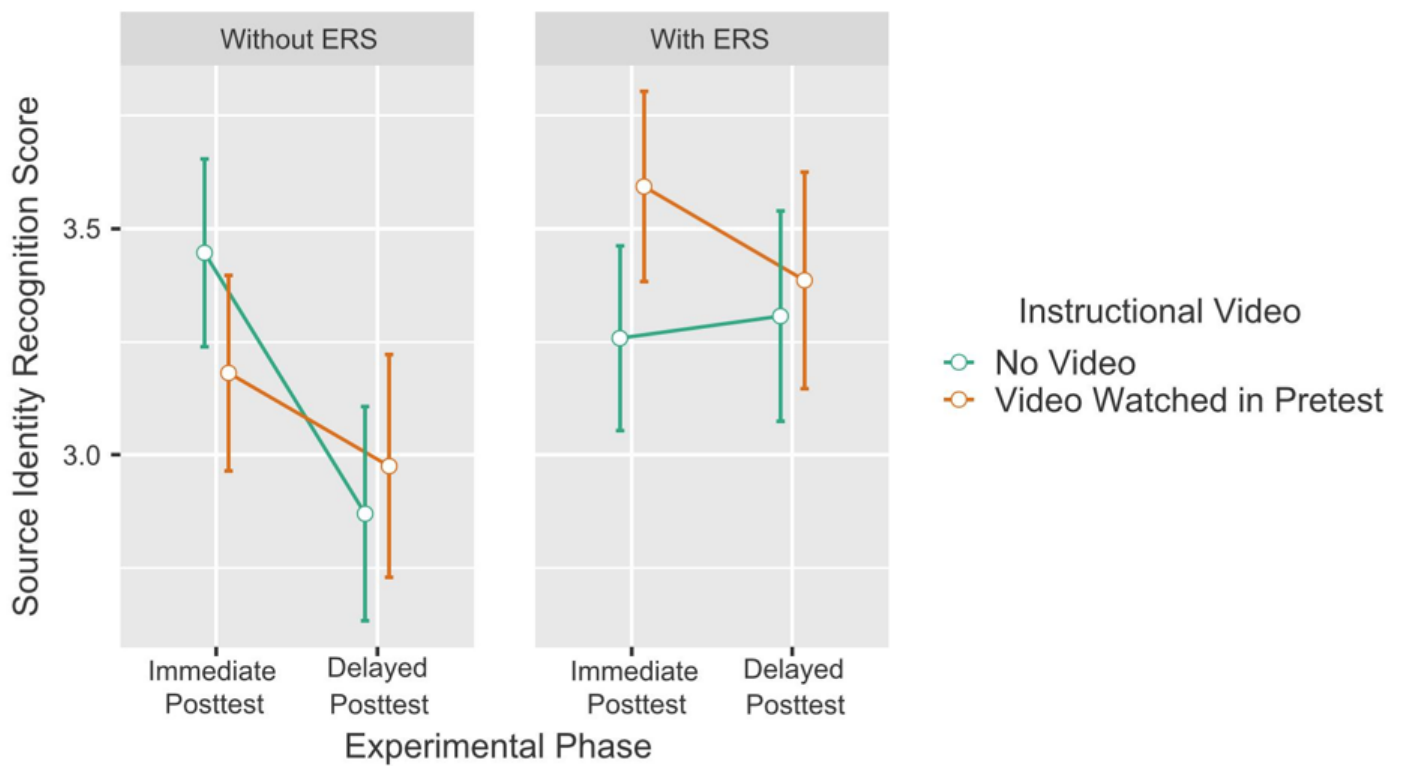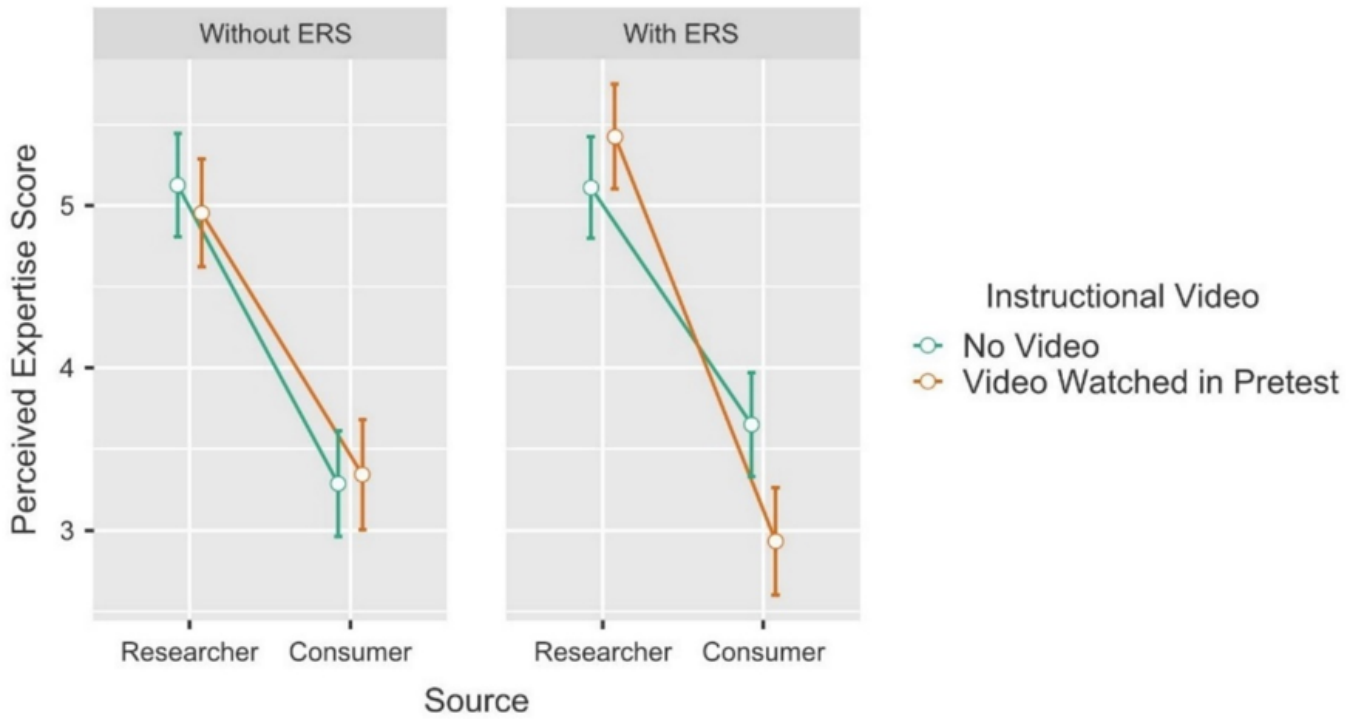**Figure 4**

Experimental Phase * Embedded Rating Scales

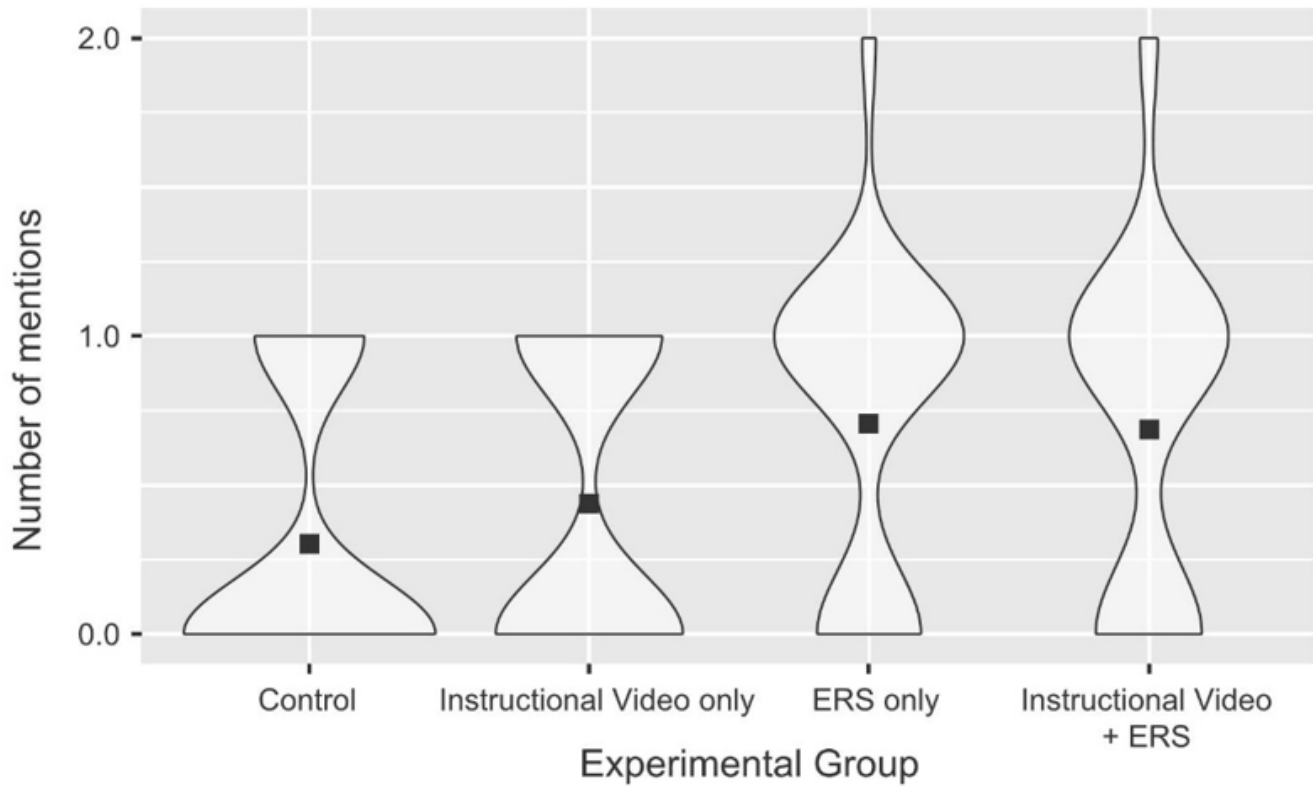## Figure 5

Experimental Phase * Embedded Rating Scales * Instructional Video

Figure 6

Source * Embedded Rating Scales * Instructional Video

Figure 7

Number of mentions of interviewees' expertise in students' justifications, according to the experimental groups

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- AppendixA.docx
- AppendixB.docx