

## Journal of Computer Assisted Learning

### Using Eye-Movement Modeling Examples to Improve Critical Reading of Multiple Webpages on a Conflicting Topic

Journal:	<i>Journal of Computer Assisted Learning</i>
Manuscript ID	Draft
Manuscript Type:	Article
Technology and Tools:	Hypertext/Hypermedia
Learning process / Pedagogy:	Individual learning
Paradigm:	Cognitivism
Level of education :	Undergraduate
Place of learning:	Formal Learning
Type of research:	Quantitative
Research technique:	Experimental
Analysis/evaluation paradigm:	PLS Regression
Issues:	Knowledge

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

Running head: EMMEs and critical reading

**Using Eye-Movement Modeling Examples to Improve Critical Reading of Multiple  
Webpages on a Conflicting Topic**

## Abstract

Interventions to promote students' source evaluations have used various methods designed for the classroom context. In the present study, we tested an approach that is easily adaptable to online courses, based on Eye Movement Modeling Examples (EMME), that is, short videos displaying an expert student's eye gaze while s/he reads multiple pages on the Internet to learn about a conflicting topic. Using an eye-tracking methodology in a pre-post design, we analyzed how EMME changed students' attention to source information, and how this processing affected their learning. EMME increased participants' attention to the search engine results page (SERP), author information, and texts from trustworthy pages, but it did not directly improve essay quality. Increases in SERP processing times mediated the positive effect of EMME on essay quality. Finally, we discuss the potential benefits and limitations of EMME in teaching complex literacy strategies, and the importance of measuring processing data in educational research studies.

**Keywords:** eye-movement modeling examples, sourcing, eye-tracking, multiple-documents comprehension.

## Using Eye-Movement Modeling Examples to Improve Critical Reading of Multiple Webpages on a Conflicting Topic

### Introduction

Digital technology has produced an increase in written communication. For years, the Internet has been the most widely used source of information. Teenagers and adults of all ages spend an increasing amount of time reading texts on the Internet for various reasons. The large number of easily available texts puts great pressure on the individual's reading skills, well beyond simple word decoding and literal comprehension. Due to the lack of information "gatekeeping" on the Internet, users can often access misinformation or biased information, which can spread quickly and become "viral". Thus, more than ever, in the so called "post-truth" era of "alternative facts", critical reading becomes crucial (Bråten, Braasch, & Salmerón, 2019).

First, critical readers must pay attention to source characteristics, such as the author's credentials or the institution publishing the document. Second, they must evaluate the quality of the information by judging to what extent these source characteristics suggest that the information is supported by sound evidence, or on the contrary, could potentially be biased. Finally, critical readers must consider the conclusions from their assessment of the source to qualify the information accessed, e.g. by judging the view supported by expert sources as credible, or by discarding information from potentially biased sources (Tarchi & Mason, 2019). Because this is a sequential process, failing to attend to source characteristics may limit students' ability to critically read the information. Accordingly, educational interventions to improve critical reading tend to emphasize the importance of paying attention to source characteristics (Brante, & Strømsø, 2017). In their recent systematic review, Brante &

1  
2  
3 Strømsø (2017) identified two relevant gaps in the literature of source evaluation  
4 interventions. First, few intervention studies were designed specifically to guide  
5 students to critically read on the Internet. As this scenario presents different source  
6 features than traditional paper documents, such as search engine result page (SERP),  
7 participants may need specific guidance to find and use such source features. Second,  
8 none of the existing interventions measured how students look for source features or  
9 specifically which features they attend to during reading. As such, the effectiveness of  
10 current interventions only relies on participants' off-line responses.  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20

21 The present study tries to fill in those gaps by testing Eye Movement Modeling  
22 Examples (EMME) to support students' source evaluations on the Internet, while  
23 measuring changes on their visual attention to source characteristics from pre to post  
24 test. EMMEs are videos that display, by means of a moving dot, where an expert is  
25 looking while performing an activity. By visually modeling expert behavior, EMMEs  
26 are intended to guide students' visual attention while performing a particular task  
27 (Jarodzka, van Gog, Dorr, Scheiter, & Gerjets, 2013; van Gog, Jarodzka, Scheiter,  
28 Gerjets, & Paas, 2009). The present study used this innovative approach to address one  
29 of the crucial questions in contemporary research on learning and instruction, that is,  
30 how to foster students' ability to read webpages critically, starting with paying attention  
31 to source information.  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45

### 46 **Sourcing and Critical Reading on the Internet**

47  
48  
49 In a seminal paper by Wineburg (1991), sourcing was defined as the process of  
50 using information about documents – author, genre, and publication date – to evaluate  
51 and interpret documents' content. Research has widely indicated that students do not  
52 tend to pay attention to source information, and this effect is generalized across different  
53 educational levels, including elementary school students (Kuiper, Wolman, & Terwel,  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 2008; Macedo-Rouet et al., 2013; Paul, Cerdán, Rouet, & Stadtler, 2018; Paul, Stadtler,  
4 & Bromme, 2019), middle school students (Mason, Ariasi, & Boldrin, 2011; Salmerón,  
5 Macedo-Rouet & Rouet, 2016; Salmerón, Agnese & Delgado, 2019; VanSledright &  
6 Kelly, 1998), high school students (Barzilai & Zohar, 2012; Macedo-Rouet et al., 2019;  
7 Paul, Macedo-Rouet, Rouet, & Stadtler, 2017; Walraven, Brand-Gruwel, & Boshuizen,  
8 2009), and university students (Barzilai, Tzadok, & Eshet-Alkalai, 2015; Bråten,  
9 Strømsø, & Andreassen, 2016; Britt & Aglinskas, 2002; List, Du, Wang, & Lee, 2019;  
10 Salmerón, Gil, & Bråten, 2018b).

11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21 Undoubtedly, sourcing is an important part of critical reading in today's  
22 information society. Readers who are not able to identify and use source information to  
23 distinguish between trustworthy and untrustworthy sources also experience more  
24 difficulties in comprehending multiple conflicting texts on the same topic (e.g., Bråten,  
25 Ferguson, Strømsø, & Anmarkrud, 2014). In fact, sourcing skills are crucial not only to  
26 be able to follow reliable sources and rule out biased ones, but also to form an  
27 integrated and coherent representation of various texts despite their divergences.  
28 Research has indicated that memory for sources is associated with the construction of  
29 coherent representations of documents' contents (Salmerón et al., 2018b; Strømsø,  
30 Bråten, & Britt, 2010).

31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45 According to the seminal documents model framework (Perfetti, Rouet, & Britt,  
46 1999; Rouet & Britt, 2011), the mental representation of a set of texts requires both the  
47 comprehension of each single text, that is, the situational model of each single text  
48 based on the integration of textual information and reader's prior knowledge (Kintsch,  
49 1989), and two additional text-representation layers, the intertext and the integrated  
50 mental models of the various situations. The intertext model is formed when readers  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 “tag” content to source characteristics such as author and genre. It refers to a  
4  
5 representation of the sources of the multiple documents and their interrelationships.  
6  
7

8         The importance of tagging content and source features also emerged in the  
9  
10 recent Discrepancy-Induced Source Comprehension (D-ISC) model (Braasch & Bråten,  
11  
12 2017). According to this model, readers pay more attention to “metadata” (who says  
13  
14 what) when they are confronted with contrasting information on the same topic  
15  
16 provided by different sources. The representation of source features also promotes the  
17  
18 interpretation of documents’ contents, contributing to the construction of an integrated  
19  
20 mental representation of multiple documents. For example, by identifying that two  
21  
22 authors have diverging motivations (e.g. to inform vs. to sell a product), readers can  
23  
24 reconcile conflicting accounts about a controversial issue.  
25  
26  
27

28         On the Internet, source features are available not only in the webpages (e.g.  
29  
30 authors’ credentials, logo of the Institution that hosts the webpage), but also in SERPs.  
31  
32 SERPs display the title of the pages, a brief summary, and the URL, usually including  
33  
34 the source that hosts the webpage. As such, they represent a unique opportunity for  
35  
36 students to reflect on the relevance and quality of the pages and directly compare pages.  
37  
38 In their recent study, Wineburg & McGrew (2017) emphasized the importance of  
39  
40 attending to SERPs in learning about conflicting topics. In their study, professional fact  
41  
42 checkers tend to spend some time scanning the SERP before clicking on any link, a  
43  
44 strategy that the authors coined as “click restraint”. They inspected the SERP to gather  
45  
46 source information from different pages, and used this in their subsequent navigation.  
47  
48  
49 On the contrary, the group of undergraduate students tended to click on the first links of  
50  
51 the list without spending much time on the SERP (for similar results see Brand-Gruwel,  
52  
53 Kammerer, van Meeuwen, & van Gog, 2017). The importance of SERPs has also been  
54  
55 identified in studies that manipulated the design to support source evaluation. For  
56  
57  
58  
59  
60

1  
2  
3 example, adapting the display of snippets to facilitate comparisons of webpages, such as  
4  
5 presenting groups of snippets instead of a list, increased the probability of students  
6  
7 including more trustworthy information in their essays (Kammerer & Gerjets, 2014) and  
8  
9 receiving higher scores on an inter-textual comprehension measure (Salmerón et al.,  
10  
11 2010). Thus, interventions aimed to improve critical reading on the Internet could  
12  
13 benefit from directly addressing how to process SERP snippets.  
14  
15

16  
17 To advance our knowledge about effective methods for promoting sourcing in  
18  
19 critical reading, in the current study we adopted the approach based on EMME, aimed at  
20  
21 modeling several strategic sourcing behaviors on the Internet: from carefully  
22  
23 scrutinizing a SERP before selecting a trustworthy source to quickly skimming a visibly  
24  
25 untrustworthy webpage.  
26  
27

### 28 **Eye Movement Modeling Example**

29  
30 EMME combines the efficacy of video-based modeling and example-based  
31  
32 instruction (Renkl, 2014) with the usefulness of an eye-tracking technology. Rooted in  
33  
34 observational learning (Bandura, 1977), video-based modeling has received increasing  
35  
36 attention in educational research as a way to foster students' performance in diverse  
37  
38 domains, such as making collages and writing poetry (Groenendijk, Janssen,  
39  
40 Rijlaarsdam, & van den Bergh, 2013), solving problems about electric circuits  
41  
42 (Hoogerheide, Renkl, Fiorella, Paas, & van Gog, 2019), or solving information  
43  
44 problems when searching the Web for information (Frerejean, van Strien, Kirschner, &  
45  
46 Brand-Gruwel, 2018). Examples are included in video-based modeling to help students  
47  
48 focus on an appropriate way to execute a task (Hefter, ten Hagen, Krense, Berthold, &  
49  
50 Renkl, 2019). Thus, they should not try to perform a task by themselves, but rather they  
51  
52 can invest mental effort in achieving optimal performance based on the example,  
53  
54 assuming that it is well-designed.  
55  
56  
57  
58  
59  
60



1  
2  
3 Eye-tracking technology provides videos of gaze replays. Therefore, it is not  
4 only useful for offering unique data about perceptual and cognitive processes, but also  
5 for providing the opportunity to show recorded visual behavior in the form of a video  
6 where fixations on specific information are represented as solid dots.  
7  
8  
9  
10

11  
12 In the extant literature, there is now ample evidence of the effectiveness of  
13 EMME in various areas of investigation, such as classification tasks (Jarodzka et al.,  
14 2012), problem solving (van Marlen, Wermeskerken, Jarodzka, & van Gog, 2018),  
15 medical imaging (Litchfield, Ball, Donovan, Manning, & Crawford,  
16 2010), medical education (Jarodzka et al., 2012; Seppänen & Gegenfurtner, 2017),  
17 digital reading (Salmerón & Llorens, 2019), and multimedia learning (Mason, Pluchino,  
18 & Tornatora, 2015, 2016; Mason, Scheiter, & Tornatora, 2017). EMME has been found  
19 to be beneficial for both the processing of information – using either texts, pictures, or  
20 medical images – and learning from it, that is, in terms of both processes and outcomes  
21 related to this type of modeling.  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33

34  
35 The mechanism that explains the effectiveness of this type of video-modeling is  
36 attention guidance. EMME guides learners' attention to what the model is focusing on  
37 at that moment, and so the visual attention of the learner is guided by and synchronized  
38 with the model, as in a state of joint attention (van Marlen et al., 2018). This attention to  
39 relevant information or aspects of the performed task contributes to deeper processing  
40 during learning and better learning outcomes. For example, in Mason and colleagues'  
41 (2015, 2016, 2017) studies, students were shown an EMME of an expert who integrated  
42 text and picture while reading an illustrated text. Findings revealed that students using  
43 EMME, compared to a control group, showed a stronger integrative processing of  
44 verbal and graphical information during the reading of a similar text, and higher  
45 learning outcomes after reading.  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 Interventions aimed at improving students' sourcing have mostly relied on  
4 worksheets, prompts, or group discussions to sustain the complex process that is part of  
5 critical reading (Wennås Brante & Strømsø, 2017). Based on the aforementioned  
6 literature on video modeling and video modeling through the gaze replay of an expert  
7 who performs well on a given task, it is theoretically legitimate to expect that EMME  
8 can be an effective way to improve sourcing. The underlying reason is that seeing what  
9 a model is looking at has a great potential to guide novices' attention to the right  
10 location at the right time (van Gog et al., 2009). EMME provides a unique opportunity  
11 in terms of attention guidance, which is at the basis of a process like sourcing, and it is  
12 initiated by focusing on usually overlooked information such as "metadata" about  
13 source characteristics. As shown in eye-tracking studies, experts and novices differ in  
14 attention allocation; that is, compared to novices, experts pay attention to task-relevant  
15 information longer and faster, and they spend less time on task-irrelevant information  
16 (e.g., van Gog, Paas, & van Merriënboer, 2005). This difference in attention allocation  
17 may lead learners to fail to encode relevant source information while reading online  
18 digital documents (e.g., Brand-Gruwel et al., 2017). Thus, EMME represents a way to  
19 reduce this discrepancy because it shows the learner what a model is looking at while  
20 performing a task (van Marlen et al., 2018).  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43

44 The link between sourcing and comprehension of multiple documents on the  
45 same topic is not only theoretically justified, as mentioned above, but it is also  
46 empirically documented in studies revealing that attention to source information is  
47 associated with better multiple-document comprehension (e.g., Anmarkrud, Bråten, &  
48 Strømsø, 2014; Salmerón et al., 2010). Therefore, if EMME supports attention to source  
49 information, we can also expect it to sustain the comprehension of conflicting  
50 information provided by various sources. It is worth noting that in some studies EMME  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 was also accompanied by verbal instructions, either unambiguous (e.g., van Gog et al.,  
4 2009) or ambiguous (van Marlen et al., 2019), to guide learners' attention, whereas in  
5  
6 others, only EMME was provided, without any verbal supplement (Mason et al., 2015,  
7  
8 2016, 2017). Based on the latter, in the current study EMME was shown without any  
9  
10 simultaneous verbal accompaniment in order to more clearly test its effectiveness in  
11  
12 modeling a complex activity like sourcing. In this way, any benefit can only be  
13  
14 attributed to the modeling itself. Moreover, redundancy between video and verbal  
15  
16 accompaniment is avoided because it can be unfavorable to learning (van Gog et al.,  
17  
18 2009). Furthermore, participants are not asked to attend to both visual and verbal  
19  
20 processes at the same time, which can be difficult. In this regard, unlike other studies  
21  
22 that used verbal accompaniments, in the present study, the sourcing task represented in  
23  
24 the EMME video did not involve visual search of small relevant elements, which may  
25  
26 not be perceived without verbal comments (Jarodzka et al., 2013).  
27  
28  
29  
30  
31  
32

33  
34 Nevertheless, sourcing processes during Internet reading are complex because  
35  
36 they involve paying attention to source information distributed across different sections  
37  
38 (e.g., snippets in the SERP, author information on a webpage), they attend longer to  
39  
40 information from a trustworthy source than from an untrustworthy source, and they use  
41  
42 information from trustworthy sources to form an integrated representation of the topic.  
43  
44 Thus, it is an open issue whether EMME, as an essentially perceptual tool, can also  
45  
46 model a process that is far from perceptual but includes some perceptual aspects.  
47  
48 However, findings on the effectiveness of EMME in modeling other non-perceptual  
49  
50 processes, such as the integration of verbal and graphical information while reading an  
51  
52 illustrated text without any verbal explanation, are encouraging (Mason et al., 2015,  
53  
54 2016, 2017).  
55  
56  
57

### 58 **The Study: Research Questions and Hypotheses**

59  
60

1  
2  
3           Sourcing is crucial when searching for information on the Internet and learning  
4  
5 in online environments. The ability to identify and represent source information  
6  
7 (metadata) to interpret a document's content and judge its authoritativeness, or to use  
8  
9 source information in referring to a document's content, is essential in order to rule out  
10  
11 or augment the content of messages on the basis of source credibility, that is, the  
12  
13 trustworthiness and accuracy of the content (Goldman & Scardamalia, 2013; Salmerón  
14  
15 et al., 2018). Given the importance of this ability in our information-saturated society,  
16  
17 this study sought to extend current research on how to improve sourcing by adopting the  
18  
19 EMME approach, which, to the best of our knowledge, has not been used in this area of  
20  
21 research. The following research questions (RQ) guided the study:  
22  
23

24           (1) Would EMME increase the time spent reading (a) the SERP snippets, (b)  
25  
26 webpage headers (i.e., the logo and the name of the institution hosting the webpage), (c)  
27  
28 information about the text author within webpages (i.e., author's name and occupation),  
29  
30 and (d) webpage texts?  
31  
32

33           (2) Would EMME increase (a) source citations and (b) ideas from trustworthy  
34  
35 pages in post-test essays?  
36  
37

38           (3) Would the effect of EMME on post-test essays be mediated by the increased  
39  
40 time spent on source information?  
41  
42

43  
44           Based on the previously reviewed literature showing that EMME has positive  
45  
46 effects on the processing of similar material by modeled learners (e.g., Jarodzka et al.,  
47  
48 2013; Mason, 2015, 2016, 2017), for RQ1 we hypothesized that the EMME group  
49  
50 would show more strategic processing of the online materials from pre to post-test, as  
51  
52 reflected in attending longer to the snippets in the SERP (RQ1a), the webpage headers  
53  
54 (RQ1b), and the authors' credentials (RQ1c). For RQ1, we also expected that students  
55  
56 in the EMME group would selectively allocate their text processing times at post-test  
57  
58  
59  
60

1  
2  
3 based on the webpage's trustworthiness, with higher reading times for trustworthy  
4 pages and lower reading times for untrustworthy pages (RQ1d).

7 For RQ2, we expected that the EMME group would cite more sources in their  
8 written essays from pre- to post-test, as a consequence of paying more attention to  
9 source information and tagging source features and contents (RQ2a) (Braasch & Bråten,  
10 2017; Bråten et al., 2014; Stang Lund, Bråten, Brandmo, Brante, & Strømsø, 2014). We  
11 also hypothesized that the EMME group would report more ideas from trustworthy  
12 pages at post-test than at pre-test, as an effect of implicit evaluation elicited by the  
13 model, who strategically looked at logos and attended to source information (RQ2b).  
14 This visual attention paid to "metadata" leads to discriminating between webpages and  
15 relying only on trustworthy online resources (Salmerón et al., 2018; Strømsø et al.,  
16 2010). The improvements in sourcing activities and products predicted for RQ 1 and 2  
17 were not expected in the control group, which received no modeling.

20 Finally, for RQ3, based on the EMME literature pertinent to the current study  
21 (Mason et al., 2015, 2016, 2017), we hypothesized that crucial processing behavior  
22 would be related to the reading outcomes. More specifically, we hypothesized that the  
23 time spent on source information at post-test would, at least partially, mediate the effect  
24 of EMME on essay quality, as indicated by an increase in source citations and ideas  
25 from trustworthy pages (Brand-Gruwel et al., 2017).

## 26 **Method**

### 27 **Participants**

28 Sixty-four undergraduate students from a large Spanish university participated in  
29 the study ( $M_{\text{age}} = 20.8$ ,  $SD = 1.91$ ; 84.1% women). Most of the students were enrolled  
30 in their third (47.6%) or fourth (38.1%) year of undergraduate psychology or education  
31 programs (52.4 and 46.0%, respectively). Students volunteered either for class credit or  
32

1  
2  
3 for an economic compensation (10€). All participants signed an informed consent form  
4  
5 and were debriefed after completing the study. From the original sample, we excluded 1  
6  
7 student due to incomplete data, which resulted in a final sample of 63 participants.  
8  
9

## 10 **Materials and Equipment**

11  
12 **Webpages.** Table 1 provides an overview of the main characteristics of the  
13  
14 webpages used. Each participant read two separate sets of four webpages on socio-  
15  
16 scientific conflicting topics. One set of webpages discussed pros and cons of the use of  
17  
18 renewable energies as a potential solution to fight climate change (CC); the other  
19  
20 discussed pros and cons of genetically modified food (GMF). The pages were  
21  
22 assembled from various authentic online texts on the issue, including institutional and  
23  
24 NGO reports, as well as diverse popular science articles. For each topic, two pages  
25  
26 provided arguments in favor of and two against the main topic (use of renewable  
27  
28 energies to fight climate change or genetically modified food). The level of  
29  
30 trustworthiness of pages was manipulated by varying the degree of expertise and  
31  
32 benevolence of the authors (Unkel & Hassel, 2017). Specifically, for pages providing a  
33  
34 similar positive or negative view, one was authored by a trustworthy source (i.e.  
35  
36 government agency or research institution) and another by an untrustworthy source (i.e.  
37  
38 a company with commercial interests in the topic or laypersons writing personal blogs).  
39  
40 Trustworthy webpages, but not untrustworthy ones, cited scientific sources in the text to  
41  
42 support their main claims.  
43  
44  
45  
46  
47  
48

49 To ensure that the webpages were appropriate for undergraduate students, we  
50  
51 computed readability scores for each webpage using the Flesch-Szigriszt Index  
52  
53 (Szigriszt, 1992), which is a version of the classic Flesch Index in Spanish. The mean  
54  
55 readability score for the CC webpages was 48.7 ( $SD = 6.1$ ), and for the GMF webpages,  
56  
57 45.8 ( $SD = 0.8$ ). According to the INFLESZ scale, values ranging between 40-55  
58  
59  
60

1  
2  
3 correspond to the category “somewhat difficult”, which includes popular science texts  
4  
5 or specialized press (Barrio-Cantalejo et al., 2008). This scale distinguishes five levels  
6  
7 of text difficulty, ranging from “very difficult” (readability < 40; e.g., undergraduate  
8  
9 textbooks) to “very easy” (readability > 80; e.g., primary school textbooks).

11  
12 [Insert Table 1 about here]

13  
14  
15 **EMME.** We constructed five EMMEs that presented a dot to represent a  
16  
17 student’s gaze on a SERP or on a particular page. Each EMME modeled a different  
18  
19 strategy corresponding to advanced readers (see Table 2). Accordingly, participants  
20  
21 were told that the EMMEs were from good students. Specifically, EMME 1 modeled a  
22  
23 student who fully inspected the results from a SERP from top to the bottom by carefully  
24  
25 reading each snippet; EMME 2 modeled a student who carefully inspected the source of  
26  
27 information provided on a webpage, including the webpage header and information  
28  
29 about the text author; EMME 3 modeled the strategy of deeply reading the text of a  
30  
31 trustworthy page by slowing down his/her gaze movements and showing re-reading  
32  
33 behavior, after attending to the source information; EMME 4 modeled the strategy of  
34  
35 skimming the text of a less trustworthy page after attending to the source information;  
36  
37 and EMME 5 modeled the strategy of quickly leaving a commercial page that was not  
38  
39 topically relevant to the student’s goal, after focusing on the source information.  
40  
41  
42  
43

44 [Insert Table 2 about here]

45  
46  
47 **Essays.** The average length of the essays was 235 words ( $SD = 65$ ,  $min = 109$ ,  
48  
49  $max = 420$ ). Responses were analyzed in terms of sourcing and comprehension. We first  
50  
51 divided each essay into ideas, defined as units with a main verb that expressed an event,  
52  
53 activity, or state (Magliano, Trabasso, & Graesser, 1999). After segmentation, essays  
54  
55 were coded to indicate whether the ideas contained an explicit reference to source  
56  
57 information, including the author’s occupation, the institution that hosted the page, and  
58  
59  
60

1  
2  
3 scientific studies referenced in the webpages. Then, we identified the webpage or pages  
4 that contained that particular idea. Finally, ideas were coded to identify students'  
5 understanding of the topic. Specifically, we distinguished among three types of ideas:  
6 single idea paraphrases, intratext inferences, and intertext inferences. Single idea  
7 paraphrases included correct claims in which students expressed an idea from one of the  
8 pages in their own words. Intratext inferences combined two single-idea paraphrases  
9 that were from one page, but not connected on the page. The paraphrase was linked to  
10 some information from students' prior knowledge. Intertext inferences combined two  
11 single-idea paraphrases from two different pages.  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22

23  
24 The first author and a trained research assistant, both blind to the conditions,  
25 independently scored a random selection of 12.7% of the essays for each of the topics: 7  
26 essays included 101 idea units for the CC topic, and 7 essays contained 92 idea units for  
27 the GMF topic. The coding of students' references to sources yielded a Cohen's Kappa  
28 of 1 for both the CC and GMF topics, and an understanding of the content of .73 and .72  
29 (for CC and GMF, respectively), thus showing substantial agreement. All disagreements  
30 were resolved through discussion between the two raters, and the research assistant  
31 scored the remaining essays according to the same coding systems.  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41

42 **Prior topic knowledge.** Prior knowledge about the topics of climate change  
43 (CC) and genetically modified food (GMF) was assessed with a true-false measure  
44 shortened and adapted from prior studies (e.g., Salmerón et al., 2010). Both measures  
45 included items about scientific and political or historical issues for each topic. The  
46 internal consistency reliability for participants' scores was questionable (Cronbach's  $\alpha =$   
47 .60 and .62 for the CC (10 items) and GMF measure (9 items), respectively).  
48  
49  
50  
51  
52  
53  
54  
55

56 **Topic interest.** Participants' personal interest in the topics of CC and GMF was  
57 assessed by means of a questionnaire that asked them about their interest and active  
58  
59  
60



1  
2  
3 involvement in the issues, using a 10-point Likert-type scale ranging from 1 (not at all  
4 true of me) to 10 (very true of me). Both measures were shortened and adapted from  
5 prior studies (e.g., Bråten, Gil, Strømsø, & Vidal-Abarca, 2009). Cronbach's  $\alpha$  for  
6 participants' answers on the CC measure (7 items) was excellent ( $\alpha = .91$ ), and it was  
7 good for answers on the GMF measure (6 items) ( $\alpha = .84$ ).  
8  
9

10  
11  
12  
13  
14  
15 **Equipment.** We used a SMI REDn eye-tracker with a sample rate of 60 Hz.  
16 BeGaze software was used to extract fixations, using the low speed event detection  
17 algorithm with a minimum threshold for fixations set at 100 msec.  
18  
19

## 20 21 **Procedure**

22  
23  
24 Participants were tested individually in a computer lab. On arrival, they were  
25 randomly assigned to the experimental or control group. First, they completed a  
26 questionnaire on demographics and filled out the prior knowledge and interest  
27 measures. Then, they were introduced to the reading and writing task. They were told to  
28 imagine that they had a personal blog and wanted to write about the topic of CC/GMF.  
29 Specifically, they were told to read the pages to write a blog entry arguing about the  
30 pros and cons of possible solutions to fight climate change (CC topic) or genetically  
31 modified food (GMF topic). After being calibrated to the eye-tracking system (using a  
32 9-point calibration), the Google SERP for the first topic was presented on the screen.  
33 From this SERP, they could access the four webpages. Task time was limited to 7  
34 minutes, but participants were also free to end the task earlier. We based this time limit  
35 on pilot testing, which indicated that seven minutes would allow all the participants to  
36 read the four webpages and re-read some of them. After they finished reading the  
37 webpages, participants were given 10 minutes to write the blog entry on a laptop using a  
38 word-processing application. Next, participants in the experimental group watched a  
39 series of EMMEs for 5 minutes. They were told that the videos displayed students' eyes  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 while performing a multiple-document task similar to what they had been doing, and  
4 that the students were chosen because they were good performers. Participants in the  
5 control condition watched a 5-minute video about the topic they had just read and  
6 written about (either CC or GMF). Finally, all the students continued with the second  
7 topic for the post-test, undergoing the same procedure as on the pre-test. The session  
8 lasted approximately 90'.  
9

## 10 11 12 13 14 15 16 17 **Design**

18  
19 We used a pre-post design with experimental and control groups. To minimize  
20 the impact of topic characteristics on the pre-post design, the presentation order was  
21 counterbalanced. In addition, to minimize the order effects from the results page, we  
22 created two SERPs with different rank orders for each topic.  
23  
24  
25  
26  
27

## 28 29 30 31 32 **Results**

### 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 **Preliminary analyses**

33 As a first step in data preprocessing, raw fixation data were examined to detect  
34 outliers, i.e. fixations of 2 SD above or below each student's mean. On average, outlier  
35 fixations represented 4.28% of individual fixations ( $SD = 0.51$ ). They were replaced by  
36 the student's median. Next, fixations were aligned with the corresponding area of  
37 interest (AOI), which included: (a) individual snippets from the SERP page, (b) the  
38 webpage hosting the institution's logo and name, corresponding to the header of each  
39 page, (c) webpage text corresponding to the paragraphs on each page, and (d) the text  
40 authors' names and affiliations included on each page in a location clearly separated  
41 from the text. Fixation times were averaged by the number of words in the  
42 corresponding AOI with textual information, in order to control for the difference in  
43 text length across webpages.  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

We also checked whether the two groups differed in their prior knowledge and interest in the two topics used in the study. An ANOVA with topic (CC and GMF) and condition (EMME vs. control) and the percentage of correct responses on the prior knowledge tests revealed non-significant effects of topic, condition, or the interaction (all  $F_s < 1$ ). A similar analysis with topic interest as dependent variable revealed a large significant effect of topic,  $F(1, 61) = 20.91, p < .001, \eta_p^2 = .26$ , and non-significant effects of condition (both  $F_s < 1$ ). Participants showed slightly higher interest in CC ( $M = 7.3, SD = 1.2$ ) than in GMF ( $M = 7.1, SD = 1.3$ ). It is important that we counterbalanced the topics across time and condition because the potential effect of topic interest was blocked in our design. More critically, prior knowledge and interest in the topic did not differ between conditions.

### **Processing of Source and Textual Information**

All the dependent variables were normally distributed, except for essay citation indices (particularly pre-test values), which were positively skewed. In these cases, logarithmic transformations were used to approximate distributions to normality.

To test the various hypotheses related to RQ1, we examined students' online processing of source information, including (a) SERP snippets, (b) webpage headers, and (c) information about the text authors. First, we performed a mixed ANOVA with condition (control and EMME) as between-participant variable, time (pre and post) as within-participant variable, and average reading time of the SERP snippets as dependent variable. Results revealed significant effects of condition,  $F(1, 61) = 5.67, p = .02, \eta_p^2 = .09$ , and time,  $F(1, 61) = 7.23, p < .01, \eta_p^2 = .11$ . These results were qualified by a significant interaction between condition and time,  $F(1, 61) = 14.75, p < .001, \eta_p^2 = .20$ . Post-hoc contrasts with Bonferroni correction indicated differences between conditions at post-test ( $p < .001$ ), but not at pre-test ( $p = .88$ ). Participants in the EMME group read

1  
2  
3 the SERP snippets longer than those in the control group only at post-test (see Table 3).  
4  
5 Across time, participants in the control group did not differ ( $p = .42$ ), whereas those in  
6  
7 the EMME group increased their SERP reading times from pre to post-test ( $p < .001$ ).  
8  
9

10 We further explored students' processing of source information within the  
11  
12 webpages. First, we performed a mixed ANOVA with condition (control and EMME)  
13  
14 as between-participant variable, time (pre- and post-test) as within-participant variable,  
15  
16 and average time spent on the webpage header as dependent variable. Results revealed a  
17  
18 significant effect of time,  $F(1, 61) = 8.04, p < .01, \eta_p^2 = .12$ , and a non-significant effect  
19  
20 of condition or the interaction (both  $F_s < 1$ ). Participants in both groups inspected the  
21  
22 headers for a longer time at post-test than at pre-test.  
23  
24  
25

26 Second, we performed a mixed ANOVA with condition (control and EMME) as  
27  
28 between-participant variable, time (pre and post-test) as within-participant variable, and  
29  
30 average time spent on the webpage's information about the text author as dependent  
31  
32 variable. Results revealed non-significant effects of time,  $F(1, 61) = 2.08, p = .16, \eta_p^2 =$   
33  
34  $.03$ , and condition ( $F < 1$ ), and a significant interaction,  $F(1, 61) = 6.08, p = .02, \eta_p^2 =$   
35  
36  $.09$ . Post-hoc analyses with Bonferroni correction indicated that, across time, the control  
37  
38 group did not differ in the time spent attending to text author information ( $p = .48$ ),  
39  
40 whereas the EMME group increased their times from pre- to post-test ( $p < .01$ ). Across  
41  
42 groups, there were no differences at pre-test ( $p = .13$ ) or at post-test ( $p = .21$ ). Whereas  
43  
44 participants in the EMME group, compared to those in the control group, spent more  
45  
46 time attending to text author information at post-test, this difference failed to reach  
47  
48 significance levels.  
49  
50  
51  
52  
53

54 Finally, for the analysis of text reading times (RQ1d), we included page  
55  
56 trustworthiness as an additional factor because in our hypothesis we predicted that the  
57  
58 experimental effects would be conditional on this factor related to source evaluation.  
59  
60

Thus, we performed a mixed ANOVA with condition (control and EMME) as between-participant variable, time (pre and post) and page trustworthiness (high or low) as within-participant variables, and average text reading time as dependent variable.

Results revealed significant effects of time,  $F(1, 61) = 8.38, p < .01, \eta_p^2 = .12$ , and page trustworthiness,  $F(1, 61) = 19.00, p < .001, \eta_p^2 = .24$ , but not condition,  $F < 1$ . There were also significant two-way interactions between time and page trustworthiness,  $F(1, 61) = 4.14, p = .04, \eta_p^2 = .06$ , and between page trustworthiness and condition,  $F(1, 61) = 4.74, p = .03, \eta_p^2 = .07$ , but not between time and condition,  $F(1, 61) = 1.97, p = .16, \eta_p^2 = .03$ . These effects were qualified by a three-way interaction between time, trust, and condition,  $F(1, 61) = 6.62, p = .01, \eta_p^2 = .10$ . To interpret this interaction, we conducted separated analyses for each time point. At pre-test, there was an effect of page trustworthiness,  $F(1, 61) = 25.67, p < .001, \eta_p^2 = .30$ , but not of condition or the interaction (both  $F_s < 1$ ). This indicates that, regardless of the condition, at pre-test participants read trustworthy pages faster than untrustworthy pages.

At post-test, there were no significant effects of page trustworthiness,  $F(1, 61) = 3.16, p = .08, \eta_p^2 = .05$ , or condition,  $F < 1$ , but there was a significant interaction between the two variables,  $F(1, 61) = 9.50, p < .01, \eta_p^2 = .14$ . Post-hoc analyses with Bonferroni correction indicated that, at post-test, participants in the EMME group, compared to those in the control group, tended to read trustworthy pages at a slower rate ( $p = .06$ ) and untrustworthy pages at a quicker rate ( $p = .03$ ). Similarly, comparisons across groups indicated that participants in the control group read trustworthy pages faster than untrustworthy pages ( $p = .001$ ), as at pre-test. This means that the control group did not change their reading approach to the texts at post-test. On the contrary, at post-test, participants in the EMME group read pages at a similar speed, regardless of their trustworthiness levels ( $p = .36$ ). This result indicated that, at post-test, participants

1  
2  
3 in the EMME group changed their reading strategy, eliminating the increased attention  
4 devoted to untrustworthy pages shown at pre-test.  
5  
6

7  
8 In sum, the results provided partial confirmation of our hypotheses related to  
9 RQ1a and RQ1c. From pre- to post-test, students in the EMME group, but not those in  
10 the control group, increased their processing time of SERP snippets and text author  
11 information within the webpages. No differences were observed in attention to webpage  
12 headers (RQ1b). Results for RQ1d, although less clear, partially support our hypothesis.  
13 Whereas participants in the control group did not change their text reading times based  
14 on the webpage's trustworthiness from pre- to post-test, those in the EMME group  
15 shifted from reading trustworthy pages faster than untrustworthy pages at pre-test to  
16 reading both types of pages at a similar rate at post-test. More critically, at post-test,  
17 participants in the EMME group showed the modeled strategies in reading texts  
18 depending on their level of trustworthiness. That is, compared to those in the control  
19 group, they tended to read trustworthy pages at a slower rate and untrustworthy pages at  
20 a quicker rate.  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36

### 37 **Source Citations and Quality of Ideas in Written Essays**

38  
39 To test the effects of EMME on source citations (RQ2), we ran a mixed  
40 ANOVA with condition (control or EMME) as between-participant variable, and time  
41 (pre and post) and source type (document and embedded) as within-participant  
42 variables. Results showed a close to significant effect of source type,  $F(1, 61) = 3.64, p$   
43  $= .06, \eta_p^2 = .06$ , indicating that participants tended to include a higher percentage of  
44 embedded sources than document sources in their essays. Non-significant effects were  
45 found for condition,  $F(1, 61) = 1.12, p = .29, \eta_p^2 = .02$ , time,  $F < 1$ , two-way interactions  
46 ( $F < 1$  for both cases), or the three-way interaction,  $F(1, 61) = 2.15, p = .15, \eta_p^2 = .03$ .  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 In addition, to test potential effects of EMME on the quality of the ideas  
4 included in the essays (RQ3), we computed a mixed ANOVA with condition (control or  
5 EMME) as between-participant variable, and time (pre and post) and idea type (single  
6 paraphrase, intra, or inter text inferences) as within-participant variables. There were  
7 non-significant main effects of condition,  $F < 1$ , time,  $F < 1$ , or idea type,  $F(1, 61) = 1.18$ ,  
8  $p = .32$ ,  $\eta_p^2 = .02$ . Similarly, we found non-significant interactions between condition  
9 and time,  $F < 1$ , condition and idea type,  $F(1, 61) = 1.81$ ,  $p = .17$ ,  $\eta_p^2 = .03$ , time and idea  
10 type,  $F < 1$ , or the three-way interaction,  $F < 1$ . Finally, we ran a mixed ANOVA with  
11 condition (control or EMME) as between-participant variable and time (pre and post) as  
12 within-participant variable on the percentage of ideas from trustworthy pages. We found  
13 non-significant effects of condition,  $F < 1$ , time,  $F < 1$ , and the interaction,  $F(1, 61) =$   
14  $2.00$ ,  $p = .16$ ,  $\eta_p^2 = .03$ .

15  
16  
17 In summary, the results did not support our hypotheses for RQ2 and RQ3  
18 because they indicate that EMME had no direct effect on the citations and quality of  
19 ideas included in the students' post-test essays.

### 20 21 22 **Relations between Reading Processes and Outcomes**

23  
24 To analyze the potential effects of EMME on the relationships between eye-  
25 movements and essay indices (RQ4), we first performed a series of Pearson correlations  
26 between all the dependent measures at post-test. Prior knowledge correlated positively  
27 with document source citations, whereas topic interest correlated negatively with the  
28 percentage of ideas from trustworthy pages included in the essays. In addition, time  
29 inspecting webpage headers and page author information correlated positively with the  
30 inclusion of document sources in the essays, but negatively with citations of embedded  
31 sources. Not surprisingly, time reading trustworthy pages correlated positively with the  
32 percentage of ideas from trustworthy pages included in the essay. Finally, reading times  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 on the SERP correlated positively with the percentage of ideas from trustworthy pages  
4 and with the inclusion of document sources in the essay.  
5  
6

7  
8 As indicated above, at post-test, participants in the EMME condition read the  
9  
10 SERP for a longer time than those in the control condition. In turn, this pattern of  
11  
12 findings allowed us to test the indirect effect of condition on sourcing in essays. In the  
13  
14 current study, we opted for the bootstrapping procedure developed by Preacher and  
15  
16 Hayes (2008), which holds no assumptions about the statistical significance of the *c*  
17  
18 path (i.e., the effect of condition on the two dependent variables tested). Specifically, we  
19  
20 tested two mediation models. First, we tested the effect of condition (EMME vs.  
21  
22 control) on the percentage of ideas from trustworthy pages included in post-test essays,  
23  
24 using the SERP reading times at post-test as a mediator. Scores for topic interest were  
25  
26 used as a covariate because they significantly correlated with the dependent variable.  
27  
28 Second, we tested a similar model with the same independent variable and mediator for  
29  
30 the document sources (i.e., the institution hosting the webpages or the texts' authors)  
31  
32 cited in post-test essays. Prior knowledge was included as covariate in the second model  
33  
34 because it correlated with the dependent variable. In both analyses, condition was  
35  
36 dummy coded (EMME = 1, control = 0), and the other variables were centered and  
37  
38 standardized. The indirect effect was tested using a bootstrap estimation approach with  
39  
40 1000 samples (Preacher & Hayes, 2008).  
41  
42  
43  
44  
45  
46

47 The first model accounted for a statistically significant portion of the variance,  
48  
49  $R^2 = .08$ ,  $F(3, 59) = 2.84$ ,  $p < .05$ . According to Cohen's (1988) effect size benchmarks,  
50  
51  $R^2$  of .08 can be considered a small effect. The bootstrapped results showed a positive  
52  
53 statistically significant indirect effect of condition on the percentage of ideas from  
54  
55 trustworthy pages included in the essays via SERP reading times, yielding an estimate  
56  
57 of 0.27 (CI<sub>95%</sub>: 0.01 - 0.88). As Figure 1 shows, the direct effect of condition on the  
58  
59  
60



percentage of ideas from trustworthy pages included in post-test essays remained statistically non-significant, which is consistent with full mediation. Finally, the covariate topic interest ( $b = -.21$ ,  $SE = .13$ ,  $p = .08$ ) was not a statistically significant predictor in this analysis.

[Insert Figures 1 and 2 about here]

Similarly, the second model resulted in a significant medium effect (Cohen, 1988),  $R^2 = .13$ ,  $F(3, 59) = 4.17$ ,  $p < .05$ . The bootstrapped results revealed a positive significant indirect effect of condition on document source citations in the post-test essays via SERP reading times, yielding an estimate of 0.27 ( $CI_{95\%}: 0.01 - 0.64$ ). As Figure 2 shows, the direct effect of condition on document source citations included in post-test essays remained non-significant, which is consistent with full mediation. Finally, the effect of the covariate prior knowledge ( $b = .21$ ,  $SE = .11$ ,  $p = .08$ ) was not significant.

To sum up, mediation analyses revealed that EMMEs exerted a positive indirect effect on the quality of students' post-test essays, as indicated by a higher percentage of ideas from trustworthy pages and a higher number of citations of document sources. Specifically, instruction with EMMEs indirectly increased approximately 27% of the standard deviation of these indicators. These effects were mediated by the increased attention given to the SERP due to the instruction with EMME.

### Discussion

This study identified, for the first time, the positive effects of a short instruction based on EMME on students' critical reading of multiple webpages about a conflicting topic. Moreover, the use of eye-tracking methods allowed us to determine how EMME affected students' processing (i.e., attention to source information) and outcomes (i.e., essay quality).

1  
2  
3 By watching a series of EMME that modeled attention to source information on  
4 SERPs and webpages, undergraduate students increased their visual attention to SERP  
5 snippets and text author information within the webpages, but not to the webpages'  
6 headers. Changes in the way the SERP was processed were particularly noteworthy  
7 (medium to high effect size), indicating that participants in the EMME group showed an  
8 increased "click restraint", moving from a novice to an expert strategy (Brand-Gruwel  
9 et al., 2017; Wineburg & McGrew, 2017). Because SERPs synthesize critical source  
10 information from all the webpages in a single space, participants could have used them  
11 to reflect on the relevance and quality of the pages and directly compare them  
12 (Kammerer & Gerjets, 2014; Salmerón et al., 2010). Corroborating this assumption, our  
13 results showed that an increased processing time of SERPs associated with the use of  
14 EMME was related to an improvement in students' essays, as indicated by increases in  
15 both the percentage of ideas from trustworthy pages and the number of citations of  
16 document sources.

17  
18  
19 The effects of EMME on students' attention to text author information within  
20 the webpages are less straightforward. EMME increased students' attention to author  
21 information, but this was not related to an improvement in their essays. More  
22 surprisingly, EMME had no impact on the processing of webpage headers. In this  
23 learning context, source information was distributed across several pages (i.e. SERPs  
24 and individual webpages), and source information provided by SERPs was partially  
25 redundant with information from the webpages (particularly the information on the  
26 headers that was already partially displayed in the snippets). The high increase in SERP  
27 processing times in the EMME group could have kept students from processing further  
28 information that was redundant. However, across all participants, visual attention to  
29 author information and headers within the webpages positively correlated with post-test  
30

1  
2  
3 essay citations of document sources, which corroborates recent findings indicating that  
4 these strategic behaviors positively support participants' inter-textual understanding of  
5 the topic (Salmerón, Gil, & Bråten, 2018a).  
6  
7  
8  
9

10 EMME also changed the way students read the texts, although this change had a  
11 limited impact on essay quality. Specifically, participants in the EMME group increased  
12 their reading times of texts from trustworthy webpages, but not from untrustworthy  
13 webpages. The control group did not change their text reading times. In addition, at  
14 post-test, the EMME group tended to devote more time to reading trustworthy pages  
15 and less time to untrustworthy pages than the control group. These effects are  
16 important, given the inherent difficulty of conveying a complex sequence of strategic  
17 decisions in EMME. Three videos from our study modeled selective reading in five  
18 steps, but only the first, second, and fifth steps were salient (carefully reading the SERP,  
19 looking at source information, and skipping a non-relevant webpage, respectively)  
20 because the third (deeply reading a trustworthy text) and fourth (skimming an  
21 untrustworthy text) were implicit. Specifically, in order to perform these implicit steps,  
22 students first had to observe that the model attended to source information. Second, they  
23 had to recognize whether the page was trustworthy or not. Third, they had to note that  
24 the model modified their reading pace (slow reading and rereading of trustworthy tests,  
25 and quick skimming of untrustworthy texts). Fourth, they had to infer that the model  
26 adjusted the reading pace to the level of trustworthiness of the page.  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48

49 In sum, participants were able to infer the strategic processing conveyed in this  
50 complex EMME, even though they were not able to benefit from this strategic behavior,  
51 as indicated by the null effects on their essays. This result opens up two different  
52 avenues for future research. First, research could analyze how students' individual  
53 differences, for example, in reading comprehension, interact with an instruction based  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 on EMME, because more advanced students could profit from them to a greater extent.  
4  
5 Second, different EMME designs can be explored, for example, by adding verbal self-  
6  
7 explanations to model the steps that cannot be conveyed visually (Salmerón & Llorens,  
8  
9 2019).

10  
11  
12 Finally, across all participants, reading times of the texts from trustworthy pages  
13  
14 correlated positively with the percentage of trustworthy ideas included in the post-test  
15  
16 essay. This result corroborates the importance of instructing students to selectively  
17  
18 attend to trustworthy information when learning about conflicting topics (García-  
19  
20 Rodicio, 2015; Goldman, Braasch, Wiley, Graesser, & Brodowinska, 2012; Wiley et al.,  
21  
22 2009).

23  
24  
25 Together, our results support the need to go beyond outcome measures in  
26  
27 educational research and identify process measures as essential components in order to  
28  
29 fully understand complex educational scenarios, such as reading multiple webpages on  
30  
31 the Internet. A common argument in favor of process measures is that they allow  
32  
33 researchers to understand how students learn, and not just what they learn (Harteis,  
34  
35 Kok, & Jarodzka, 2018). However, our study exemplifies a different and probably more  
36  
37 critical argument in favor of process data. Our instructional manipulation had no direct  
38  
39 effect on essay quality, but it did have an indirect effect via SERP processing. If we had  
40  
41 not used eye-tracking measures to analyze how participants processed source  
42  
43 information, we would have erroneously concluded that EMME had no effect on essay  
44  
45 quality. This is partly due to the fact that participants show great variability in terms of  
46  
47 the strategic processes being taught. Preexisting processing differences within and  
48  
49 between groups could hide the effects of a manipulation on the learning output if they  
50  
51 are not considered in the analyses. Thus, combining process and outcome measures in  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 statistical models, as in the path analyses used in this study, helps researchers to avoid  
4  
5 committing type II errors.  
6

### 7 **Educational Implications**

8  
9  
10 As pointed out in previous investigations, the eye movement modeling example  
11  
12 is not only a research tool but also an instructional tool because it can be used in  
13  
14 classrooms to model the use of effective strategies for successful performance on  
15  
16 complex tasks and activities (Mason et al., 2016). Once the videos are prepared, they  
17  
18 can easily be used by teachers and instructors when introducing a new strategy or  
19  
20 sequence of strategies. These videos can emphasize the importance of paying attention  
21  
22 to specific elements, explaining that through appropriate visual behavior, students  
23  
24 perform better. Teachers and instructors can also highlight that, even for complex  
25  
26 reading tasks that require far more than perceptual processes, the latter are at the basis  
27  
28 of information processing. In other words, EMME can contribute to refining students'  
29  
30 metacognitive awareness of the various steps into which a complex process can be  
31  
32 divided, and the specific strategies that are powerful in completing a task or activity in a  
33  
34 learning context.  
35  
36  
37  
38  
39

40 Paying attention to relevant elements through perceptual cues – like gaze replays  
41  
42 – for appropriate encoding of information is the first step in moving toward successful  
43  
44 performance on complex and demanding tasks (Jarodzka et al, 2013; van Marlen et al.,  
45  
46 2018). Source evaluation for critical reading is one of these tasks that require visual  
47  
48 attention to particular information in order to discriminate between trustworthy and  
49  
50 untrustworthy sources. Even short videos showing the visual behavior of a successful  
51  
52 performer seem to be effective, at least to some extent, for both the processes and  
53  
54 outcomes when reading conflicting documents on the same topic. Therefore, EMME  
55  
56 can be implemented in a relatively short time.  
57  
58  
59  
60

1  
2  
3 Moreover, EMMEs can represent a starting point of tutorials, particularly in  
4  
5 blended and online learning environments, and they can also take advantage of learning  
6  
7 at one's own time and pace (e.g., Rienties, Tempelaar, Nguyen, & Littlejohn, 2019).  
8  
9 Through attentional guidance, EMME contributes to making some essential aspects  
10  
11 salient to students in order to increase their online and offline performance, as revealed  
12  
13 by process and outcome measures. Even tasks that involve higher-order thinking  
14  
15 processes, such as critical reading, are based on appropriate encoding of relevant  
16  
17 information during effective processing of text content.  
18  
19

### 20 21 **Limitations**

22  
23 This study has some limitations. The first is that the design included only two  
24  
25 conditions. Eye-tracking studies with complex learning materials are laborious, and  
26  
27 practical constraints do not always allow optimal research designs. However, a stronger  
28  
29 investigation on the potential of EMME to enhance source evaluations should add more  
30  
31 conditions, for example, the use of a different control condition with more traditional  
32  
33 video-modeling, or a condition with EMME and the generation of observers' self-  
34  
35 explanations.  
36  
37

38  
39 The second limitation is the use of only four webpages (one text each) for each  
40  
41 topic. In the future, more complex designs should also include a set of documents from  
42  
43 sources that vary in trustworthiness to a more subtle degree, so that it becomes more  
44  
45 challenging to differentiate among the sources. Related to this, the third limitation is  
46  
47 that we assumed the trustworthiness of the four webpages on the basis of the sources'  
48  
49 characteristics, but we do not know whether the participants perceived the sources'  
50  
51 characteristics in the same way. However, we are inclined to assume that university  
52  
53 students are able to distinguish between a webpage of the World Health Organization  
54  
55 (WHO) and a webpage of a company like Monsanto, or a page of the United Nations  
56  
57  
58  
59  
60

1  
2  
3 Environment Program (UNEP) and a personal blog, when they encode source  
4  
5 information. Nevertheless, future investigations will benefit from a manipulation check  
6  
7 regarding webpages' trustworthiness.  
8  
9

10 The fourth limitation is that we considered only textual information. However,  
11  
12 when we search the Internet for information, we face multimedia materials that  
13  
14 introduce static or dynamic visualizations. It seems worthwhile to investigate attentional  
15  
16 guidance in relation to both types of representations, as well as the integration of verbal  
17  
18 and graphic information on pages with different levels of trustworthiness.  
19  
20

## 21 **Conclusions**

22  
23  
24 Despite these limitations, this study is the first to indicate that a short EMME  
25  
26 can be used to foster important aspects of source evaluation processes for critical  
27  
28 reading of documents on debated issues. Attention to the results that appear on a SERP  
29  
30 after a search, and to source information within the accessed pages, can be increased  
31  
32 through EMME. Moreover, the study also provides evidence that EMME play a role in  
33  
34 post-reading essays about the contents read online – in terms of the percentage of ideas  
35  
36 from trustworthy pages and the document sources cited – through the mediating  
37  
38 variable of the time spent on the SERP. The more time the participants spent carefully  
39  
40 reading the snippets of SERP in the EMME condition, the more they reported ideas  
41  
42 from trustworthy pages and document sources in their written essays. Thus, EMME are  
43  
44 to some extent effective in enhancing, either directly or indirectly, the processing of the  
45  
46 reading material, which then has beneficial consequences in terms of the ideas that  
47  
48 readers recall from trustworthy sources when engaging in critical reading.  
49  
50  
51  
52  
53  
54  
55

## 56 **References**

57  
58  
59  
60

- 1  
2  
3 Anmarkrud, O., Bråten, I., & Strømsø, H. I. (2014). Multiple-documents literacy:  
4  
5 Strategic processing, source awareness, and argumentation when reading  
6  
7 multiple conflicting documents. *Learning and Individual Differences*, 30, 64–76.  
8  
9 doi:10.1016/j.lindif.2013.01.007  
10  
11  
12 Bandura, A. (1977). *Social learning theory*. Englewood Cliffs, NJ: Prentice-Hall.  
13  
14 Barrio-Cantalejo, I. M., Simón-Lorda, P., Melguizo, M., Escalona, I., Marijuán, M. I., &  
15  
16 Hernando, P. (2008). Validación de la Escala INFLESZ para evaluar la  
17  
18 legibilidad de los textos dirigidos a pacientes [Validation of the INFLESZ scale  
19  
20 to evaluate the readability of patient texts]. *Anales del Sistema Sanitario de*  
21  
22 *Navarra*, 31, 135-152.  
23  
24  
25  
26 Barzilai, S., Tzadok, E., & Eshet-Alkalai, Y. (2015). Sourcing while reading divergent  
27  
28 expert accounts: Pathways from views of knowing to written argumentation.  
29  
30 *Instructional Science*, 43, 737-766. doi:10.1007/s11251-015-9359-4  
31  
32  
33 Barzilai, S., & Zohar, A. (2012). Epistemic thinking in action: Evaluating and  
34  
35 integrating online sources. *Cognition and Instruction*, 30, 39-85.  
36  
37 doi:10.1016/j.learninstruc.2010.02.008  
38  
39  
40 Braasch, J. L. G., & Bråten, I. (2017). The discrepancy-induced source comprehension  
41  
42 (D-ISC) model: Basic assumptions and preliminary evidence. *Educational*  
43  
44 *Psychologist*, 52, 167-181. doi:10.1080/00461520.2017.1323219  
45  
46  
47 Braasch, J. L. G., Bråten, I., Strømsø, H. I., Anmarkrud, O., & Ferguson, L. E. (2013).  
48  
49 Promoting secondary school students' evaluation of source features of multiple  
50  
51 documents. *Contemporary Educational Psychology*, 38, 180-195.  
52  
53 doi:10.1016/j.cedpsych.2013.03.003  
54  
55  
56  
57  
58  
59  
60



- 1  
2  
3 Brand-Gruwel, S., Kammerer, Y., van Meeuwen, L., & van Gog, T. (2017). Source  
4 evaluation of domain experts and novices during Web search. *Journal of Computer*  
5 *Assisted Learning*, 33, 234-251. doi:10.1111/jcal.12162  
6  
7  
8  
9  
10 Bråten, I., Braasch, J., & Salmerón, L. (2019). Reading multiple and non-traditional  
11 texts: new opportunities and new challenges. In E. B. Moje, P. Afflerbach, P.  
12 Enciso, & N. K. Lesaux (Eds.), *Handbook of Reading Research* (Vol. 5). New  
13 York: Routledge.  
14  
15  
16  
17  
18  
19 Bråten, I., Ferguson, L. E., Strømsø, H. I., & Anmarkrud, Ø (2014). Students working  
20 with multiple conflicting documents on a scientific issue: Relations between  
21 epistemic cognition while reading and sourcing and argumentation in essays.  
22 *British Journal of Educational Psychology*, 84, 58-85. doi:10.1111/bjep.12005  
23  
24  
25  
26  
27  
28 Bråten, I., Gil, L., Strømsø, H. I., & Vidal-Abarca, E. (2009). Personal epistemology  
29 across cultures: Exploring Norwegian and Spanish university students' epistemic  
30 beliefs about climate change. *Social Psychology of Education*, 12, 529-560.  
31 doi:10.1007/s11218-009-9097-z  
32  
33  
34  
35  
36  
37 Bråten, I., Strømsø, H. I., & Andreassen, R. (2016). Sourcing in professional education:  
38 Do text factors make any difference? *Reading and Writing*, 29, 1599-1628. doi:  
39 10.1007/s11145-015-9611-y  
40  
41  
42  
43  
44  
45 Britt, M. A., & Aglinskias, C. (2002). Improving students' ability to identify and use  
46 source information. *Cognition and Instruction*, 20, 485-522.  
47 doi:10.1207/S1532690XCI2004\_2  
48  
49  
50  
51 Delgado, P., Avila, V., Fajardo, I., & Salmerón, L. (2019). Training young adults with  
52 intellectual disability to read critically on the Internet. *Journal of Applied Research*  
53 *in Intellectual Disabilities*, 32, 666-677. doi:10.1111/jar.12562  
54  
55  
56  
57  
58  
59  
60

- 1  
2  
3 Frerejean, J., van Strien, J. L. H., Kirschner, P. A., & Brand-Gruwel, S. (2018). Effects  
4 of a modelling example for teaching information problem solving skills. *Journal of*  
5 *Computer Assisted Learning*, 34, 688-700. doi:10.1111/jcal.12276  
6  
7  
8  
9  
10 García-Rodicio, H. (2015). Students' evaluation strategies in a Web research task: Are  
11 they sensitive to relevance and reliability? *Journal of Computing in Higher*  
12 *Education*, 27, 134–157. doi: 10.1007/s12528-015-9098-1  
13  
14  
15  
16  
17 Goldman, S.R., Braasch, J.L.G., Wiley, J., Graesser, A.C., & Brodowinska, K. (2012).  
18 Comprehending and learning from Internet sources: Processing patterns of better  
19 and poorer learners. *Reading Research Quarterly*, 47, 356–381.  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60
- Frerejean, J., van Strien, J. L. H., Kirschner, P. A., & Brand-Gruwel, S. (2018). Effects of a modelling example for teaching information problem solving skills. *Journal of Computer Assisted Learning*, 34, 688-700. doi:10.1111/jcal.12276
- García-Rodicio, H. (2015). Students' evaluation strategies in a Web research task: Are they sensitive to relevance and reliability? *Journal of Computing in Higher Education*, 27, 134–157. doi: 10.1007/s12528-015-9098-1
- Goldman, S.R., Braasch, J.L.G., Wiley, J., Graesser, A.C., & Brodowinska, K. (2012). Comprehending and learning from Internet sources: Processing patterns of better and poorer learners. *Reading Research Quarterly*, 47, 356–381. doi:10.1002/RRQ.027
- Groenendijk, T., Janssen, T., Rijlaarsdam, G., & van den Bergh, H. (2013). The effect of observational learning on students' performance, processes, and motivation in two creative domains. *British Journal of Educational Psychology*, 83, 3-28. doi:10.1111/j.2044-8279.2011.02052.x
- Harteis, C., Kok, E. M., & Jarodzka, H. (2018). The journey to proficiency: Exploring new objective methodologies to capture the process of learning and professional development. *Frontline Learning Research*, 6, 1-5. doi:10.14786/flr.v6i3.435
- Hefter, M. H., ten Hagen, I. Krense, C., Berthold, K., & Renkl, A. (2019). Effective and efficient acquisition of argumentation knowledge by self-explaining examples: videos, texts, or graphic novels? *Journal of Educational Psychology*. Advance online publication. doi:10.1037/edu0000350
- Hoogerheide, V., Renkl, A., Fiorella, L., Paas, & van Gog, T. (2019). Enhancing example-based learning: Teaching on video increases arousal and improves problem-solving performance. *Journal of Educational Psychology*, 111, 45-56. doi.org/10.1037/edu0000272

- 1  
2  
3 Jarodzka, H., Balslev, T., Holmqvist, K., Nyström, M., Scheiter, K., Gerjets, P., & Eika,  
4  
5 B. (2012). Conveying clinical reasoning based on visual observation via eye-  
6  
7 movement modelling examples. *Instructional Science*, *40*, 813-827.  
8  
9 doi:10.1007/s11251-012-9218-5  
10  
11  
12 Jarodzka, H., van Gog, T., Dorr, M., Scheiter, K., & Gerjets, P. (2013). Learning to see:  
13  
14 Guiding students' attention via a Model's eye movements fosters learning.  
15  
16 *Learning and Instruction*, *25*, 62-70. doi:10.1016/j.learninstruc.2012.11.004  
17  
18  
19 Kammerer, Y., Bråten, I., Gerjets, P., & Strømsø, H. I. (2013). The role of Internet-  
20  
21 specific epistemic beliefs in laypersons source evaluations and decisions during  
22  
23 Web search on a medical source. *Computers in Human Behavior*, *29*, 1193-1203.  
24  
25 doi:10.1016/j.chb.2012.10.012  
26  
27  
28 Kammerer, Y., & Gerjets, P. (2014). The role of search result position and source  
29  
30 trustworthiness in the selection of web search results when using a list or a grid  
31  
32 interface. *International Journal of Human-Computer Interaction*, *30*, 177-191.  
33  
34 doi:10.1080/10447318.2013.846790  
35  
36  
37 Kammerer, Y, Meier, N., & Stahl, E. (2016). Fostering secondary-school students'  
38  
39 intertext model formation when reading a set of websites: The effectiveness of  
40  
41 source prompts. *Computers & Education*, *102*, 52-64.  
42  
43 doi:10.1016/j.compedu.2016.07.001  
44  
45  
46 Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. New York: Cambridge  
47  
48 University Press.  
49  
50  
51 Kuiper, E., Wolman, M., & Terwel, J. (2008). Integrating critical Web skills and  
52  
53 content knowledge: Development and evaluation of a 5th grade educational  
54  
55 program. *Computers in Human Behavior*, *24*, 666-692.  
56  
57 doi:10.1016/j.chb.2007.01.022  
58  
59  
60

- 1  
2  
3 List, A., Du, H., Wang, Y., & Lee, H. Y. (2019). Toward a typology of integration:  
4  
5 Examining the documents model framework. *Contemporary Educational*  
6  
7 *Psychology*, 58, 228-242. doi:10.1016/j.cedpsych.2019.03.003  
8  
9
- 10 Litchfield, D., Ball, L. J., Donovan, T., Manning, D. J., & Crawford, T. (2010). Viewing  
11  
12 another person's eye movements improves identification of pulmonary nodules  
13  
14 in chest X-ray inspection. *Journal of Experimental Psychology: Applied*, 16,  
15  
16 251-262. doi:10.1037/a0020082  
17  
18
- 19 Macedo-Rouet, M., Braasch, J. L. G., Britt, A. M., & Rouet, J-F. (2013). Teaching  
20  
21 fourth and fifth graders to evaluate information sources during text  
22  
23 comprehension. *Cognition and Instruction*, 31, 204-226.  
24  
25 doi:10.1080/07370008.2013.769995  
26  
27
- 28 Macedo-Rouet, M., Potocki, A., Scharrer, L., Ros, C., Stadtler, M., Salmerón, & Rouet,  
29  
30 J-F. (2019). *How good is this page? Benefits and limits of prompting on*  
31  
32 *adolescents' evaluation of web information quality*. Advance online publication.  
33  
34 doi:10.1002/rrq.241  
35  
36
- 37 Magliano, J. P., Trabasso, T., & Graesser, A. C. (1999). Strategic processes during  
38  
39 comprehension. *Journal of Educational Psychology*, 91, 615-629.  
40  
41 doi:10.1037/0022-0663.91.4.615  
42  
43
- 44 Mason, L., Ariasi, N., & Boldrin, A. (2011). Epistemic beliefs in action: Spontaneous  
45  
46 reflections about knowledge and knowing during online information searching  
47  
48 and their influence on learning. *Learning and Instruction*, 21, 137-151.  
49  
50 doi:10.1016/j.learninstruc.2010.01.001  
51  
52
- 53 Mason, L., Pluchino, P., & Tornatora, M. C. (2015). Eye-movement modeling of text  
54  
55 and picture integration during reading: effects on processing and learning.  
56  
57  
58  
59  
60

1  
2  
3 *Contemporary Educational Psychology*, 41, 172-187.

4  
5 doi:10.1016/j.cedpsych.2015.01.004

6  
7  
8 Mason, L., Pluchino, P., & Tornatora, M. C. (2016). Using eye-tracking technology as  
9  
10 an indirect instruction tool to improve text and picture processing and learning.  
11  
12 *British Journal of Educational Technology*, 47, 1083-1095.

13  
14 doi:10.1111/bjet.12271

15  
16  
17 Mason, L., Scheiter, K., & Tornatora, M. C. (2017). Using eye movements to model the  
18  
19 sequence of text-picture processing for multimedia comprehension *Journal of*  
20  
21 *Computer Assisted Learning*, 33, 443-460. doi:10.1111/jcal.12191

22  
23  
24 Paul, J., Cerdán, R., Rouet, J. F., & Stadtler, M. (2018). Exploring fourth graders'  
25  
26 sourcing skills/Un análisis de la capacidad de escrutinio sobre las fuentes de  
27  
28 información de los estudiantes de cuarto grado. *Infancia y Aprendizaje*, 41, 536-  
29  
30 580. doi:10.1080/02103702.2018.1480458

31  
32  
33 Paul, J., Stadtler, M., & Bromme, R. (2019). Effects of a sourcing prompt and conflicts  
34  
35 in reading materials on elementary students' use of source information.  
36  
37 *Discourse Processes*, 56, 155-169. doi:10.1080/0163853X.2017.1402165

38  
39  
40 Preacher, K.J., & Hayes, A.F. (2008). Asymptotic and resampling strategies for  
41  
42 assessing and comparing indirect effects in multiple mediator models. *Behavior*  
43  
44 *Research Methods*, 40, 879-891. doi:10.3758/BRM.40.3.879

45  
46  
47 Renkl, A. (2014). Toward an instructionally oriented theory of example-based learning.  
48  
49 *Cognitive Science*, 38, 1-37. doi:10.1111/cogs.12086

50  
51  
52 Rienties, B., Tempelaar, D., Nguyen, Q., & Littlejohn, A. (2019). Unpacking the  
53  
54 intertemporal impact of self-regulation in a blended mathematics environment.  
55  
56 *Computers in Human Behavior*. Advance online publication.

57  
58 doi:10.1016/j.chb.2019.07.007  
59  
60

- 1  
2  
3 Salmerón, L., Gil, L., & Bråten, I. (2018a). Using eye-tracking to assess sourcing during  
4 multiple document reading: A critical analysis. *Frontline Learning Research*,  
5 6, 105-122. doi:10.14786/flr.v6i3.368  
6  
7  
8  
9
- 10 Salmerón, L., Gil, L., & Bråten, I. (2018b). Effects of reading real versus print-out  
11 versions of multiple documents on students' sourcing and integrated  
12 understanding. *Contemporary Educational Psychology*, 52, 25-35.  
13  
14  
15  
16  
17  
18  
19  
20  
21 Salmerón, L., Gil, L., Bråten, I., & Strømsø, H. (2010). Comprehension effects of  
22 signalling relationships between documents in search engines. *Computers in*  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60
- Salmerón, L., & Llorens, A. (2019). Instruction of digital reading strategies based on  
eye-movements modeling examples. *Journal of Educational Computing Research*,  
57, 343-359. doi:10.1177/0735633117751605
- 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. doi:10.1002/asi.23585
- Salmerón, L., Strømsø, H. I., Kammerer, Y., Stadtler, M., & van den Broek, P. (2019).  
Comprehension processes in digital reading. In M. Barzillai, J. Thomson, S.  
Schroeder, & P. van den Broek (Eds.). *Learning to read in a digital world (pp. 91-  
120)*. Amsterdam, John Benjamins.
- Seppänen, M., & Gegenfurtner, A. (2012). Seeing through a teacher's eye improves  
students' imagining interpretation. *Medical Education*, 46, 1113-1114.  
doi:10.1111/medu.12041

- 1  
2  
3 Szigriszt, F. (1992). *Sistemas predictivos de legibilidad del mensaje escrito: Fórmula*  
4  
5 *de perspicuidad*. Unpublished doctoral dissertation, Universidad Complutense  
6  
7 de Madrid.  
8  
9
- 10 Strømsø, H. I., Bräen, I., & Britt, A. M. (2010). Reading multiple texts about climate  
11  
12 change: The relationship between memory for sources and text comprehension.  
13  
14 *Learning and Instruction, 20*, 192-204. doi:10.1016/j.learninstruc.2009.02.001  
15  
16
- 17 Tarchi, C., & Mason, L. (2019). Effects of critical thinking on multiple-document  
18  
19 comprehension. *European Journal of Psychology of Education*. Advance online  
20  
21 publication. doi:10.1007/s10212-019-00426-8  
22  
23
- 24 van Gog, T., Jarodzka, H., Scheiter, K., Gerjets, P., & Paas, F. (2009). Attention  
25  
26 guidance during example study via the model's eye movements. *Computers in*  
27  
28 *Human Behavior, 25*, 785–791. doi:10.1016/j.chb.2009.02.007  
29  
30
- 31 van Gog, T., Paas, F., & Van Merriënboer, J. J. G. (2005). Uncovering expertise-  
32  
33 related differences in troubleshooting performance: Combining eye movement  
34  
35 and con-current verbal protocol data. *Applied Cognitive Psychology, 19*, 205–  
36  
37 221. <https://doi.org/10.1002/acp.1112>  
38  
39
- 40 van Marlen, T., Wermeskerken, M., Jarodzka, H., & van Gog, T. (2018). Effectiveness  
41  
42 of eye movement modeling examples in problem solving: The role of verbal  
43  
44 ambiguity and prior knowledge. *Learning and Instruction, 58*, 274-283.  
45  
46 doi:10.1016/j.learninstruc.2018.07.005  
47  
48
- 49 VanSledright, B., & Kelly, C. (1998). Reading American history: The influence on  
50  
51 multiple sources on six fifth graders. *The Elementary School Journal, 98*, 239-  
52  
53 265.  
54  
55  
56  
57  
58  
59  
60

- 1  
2  
3 Brante, E., Strømsø, H. I. (2017). Sourcing in text comprehension: a review of  
4  
5 interventions targeting sourcing skills. *Educational Psychology Review*, 30,  
6  
7 773–799. doi:10.1007/s10648-017-9421-7  
8  
9
- 10 Wiley, J., Goldman, S., Graesser, A., Sanchez, C., Ash, I. & Hemmerich, J. (2009).  
11  
12 Source evaluation, comprehension, and learning in internet science inquiry tasks.  
13  
14 *American Educational Research Journal*, 46, 1060-1106.  
15  
16 doi:10.3102%2F0002831209333183  
17  
18
- 19 Wineburg, S. (1991). Historical problem solving: A study of the cognitive processes  
20  
21 used in the evaluation of documentary and pictorial evidence. *Journal of*  
22  
23 *Educational Psychology*, 83, 73-87. doi:10.1037/0022-0663.83.1.73  
24  
25
- 26 Wineburg, S. & McGrew, S. (2017). Lateral reading: Reading less and learning more  
27  
28 when evaluating digital information. *Stanford History Education Group Working*  
29  
30 *Paper No. 2017-A1*. doi:10.2139/ssrn.3048994  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60



Table 1

*Overview of the webpages for the two topics.*

Topic	Webpage	Author	Content	Number of words
	World Health Organization (WHO)	Director of Environmental Health of WHO	Argues in favor of GMF. States that GMF contain more vitamins and minerals and can be better conserved. The text cites a scientific study to support this view.	272
GMF	Monsanto (Global Agriculture Company)	Director of department of risk management from Monsanto	Argues in favor of GMF. Outlines higher resistance to insects or diseases of GM crops and higher productivity.	271
	Food and Agriculture Organization of the United Nations (FAO)	Spanish representative of FAO	Argues against GMF. Discusses ecological risks of GM crops for surrounding soil and plants. The text cites a scientific study to support this view.	265
	Personal blog	Economist	Argues against GMF. State potential toxicity of GM proteins.	279
	Iberdrola (Spanish multinational electric utility company)	Iberdrola's project director	The company sector addressed with renewable energy business. Argues in favor of solutions to CC. States the advantages of wind, solar, and hydroelectric power.	252
CC	United Nations Environment Program (UNEP)	Director of UNEP	Argues in favor of using renewable energies as a solution to CC. Discusses reduction of carbon dioxide emissions by using biodiesel. The text cites a scientific study to support this view.	271
	Personal blog	Undergraduate student in History	Argues against using renewable energies. Outlines that biofuels make CC worse.	269
	I2C2 (Spanish climate change research institute)	Communication director of Research Institute	Argues against using renewable energies. Discusses great economic costs and little success of renewable energies. The text cites a scientist from Harvard university to support the claims raised.	296

*Note.* GMF = Genetically Modified Food; CC = Climate Change

Table 2

*Overview of the EMMEs used.*






EMME #	Strategy modeled	Description	Screenshot	Time
1	Full SERP inspection	A student inspects a SERP page from top to bottom, reading at a normal pace all the page titles and some further information from the snippets. The student ends up clicking on a relevant page at the bottom of SERP after a review of the SERP titles.		69''
2	Identification of source information	A student looks at the webpage logo, reads the text once at normal pace, and finally reads the author information provided below the text.		52''
3	Deep reading of trustworthy and relevant pages	A student looks at the webpage logo (institutional page), reads the text twice at normal pace, and finally reads the author information provided below the text.		77''
4	Skim less trustworthy and irrelevant pages	A student looks at the webpage logo (popular forum) and user's information located at the left of the text, and quickly skims the text.		29''
5	Quickly abandon topically unrelated pages	A student looks at the webpage logo (commercial service unrelated to the task) and abandons the page without reading the text.		18''

Table 3

*Means and standard deviation (in brackets) for measured variables.*

	Control		EMME	
	Pre-test	Post-test	Pre-test	Post-test
<i>Eye-movements indices</i>				
SERP reading times	145.2 (84.5)	132.3 (59.7)	148.4 (84.9)	221.9 (117.4)
Page header dwell times	700.2 (418.3)	1072.6 (772.2)	789.9 (467.1)	1079.9 (834.1)
Page author information dwell times	1486.4 (1122.6)	1342.0 (886.4)	1105.0 (800.7)	1656.0 (1106.0)
Trustworthy pages reading times	267.5 (45.8)	258.3 (36.6)	271.1 (55.7)	285.1 (55.7)
Untrustworthy pages reading times	312.7 (61.2)	310.6 (57.2)	317.7 (65.3)	271.1 (80.6)
<i>Essay indices</i>				
Single paraphrases	0.7 (1.0)	0.8 (1.5)	0.7 (1.0)	0.8 (1.5)
Intra-text inference	0.7 (0.8)	0.8 (1.0)	1.1 (1.1)	1.0 (0.9)
Inter-text inference	1.1 (0.9)	1.1 (0.8)	0.8 (0.8)	0.9 (0.8)
Percentage of ideas from trustworthy pages	14.5 (9.7)	13.9 (6.4)	13.4 (8.0)	16.2 (10.8)
Document sources	0.3 (0.6)	0.3 (0.7)	0.5 (1.1)	0.7 (0.9)
Embedded sources	0.5 (1.0)	0.8 (1.0)	0.7 (0.9)	0.6 (0.8)

Table 4

*Zero-order correlations between condition and measured variables at post-test*

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Condition	-													
2. Prior knowledge	.08	-												
3. Topic interest	-.09	.06	-											
4. SERP reading times	.44**	.13	.03	-										
5. Page header dwell times	.01	.24	.11	.44**	-									
6. Page author information dwell times	.16	.13	-.01	.43**	.66**	-								
7. Trustworthy pages reading times	.24	-.19	-.03	.09	.14	.12	-							
8. Untrustworthy pages reading times	-.28*	-.25*	-.06	-.22	-.24	-.22	.03	-						
9. Single idea paraphrase	.02	-.22	-.12	-.13	-.10	-.03	.12	.21	-					
10. Intra-text inference	.14	-.10	-.19	.12	-.21	-.10	.03	.14	.16	-				
11. Inter-text inference	-.12	.04	.02	-.03	.09	.03	.08	.07	-.24	-.50**	-			
12. % ideas from trustworthy pages	.13	-.07	-.27*	.26*	-.05	-.08	.25*	.01	.21	.15	.24	-		
13. Document source citations	.23	.26*	-.09	.35*	.28*	.30*	-.09	-.05	-.27*	-.04	.32*	.17	-	
14. Embedded source citations	-.12	-.05	-.04	-.08	-.27*	-.22	.02	.17	.21	.01	.35**	.20	-.09	-

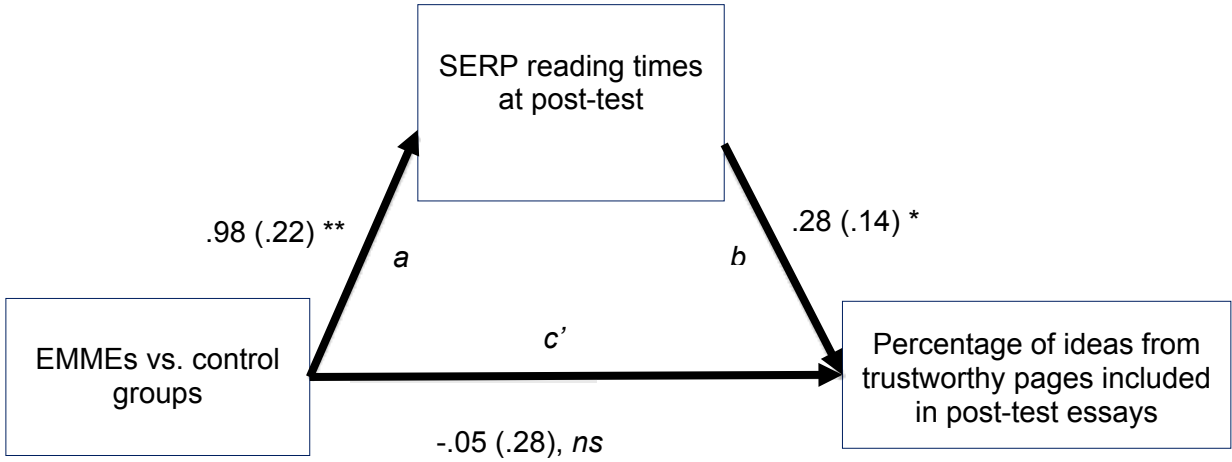
Note. \* $p < .05$ , \*\* $p < .01$ .

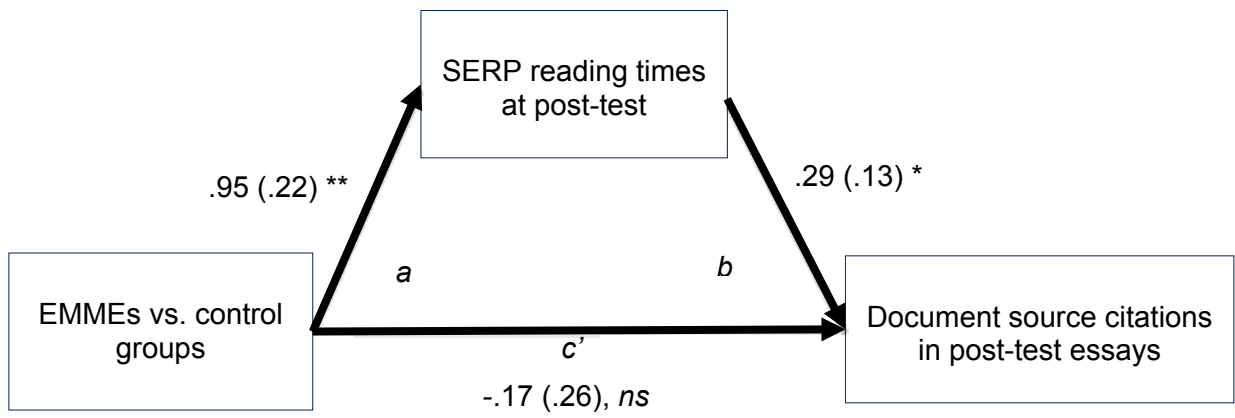
## Figure Caption

**Figure 1.** Mediation model for the indirect effect of condition (dummy coded 1 = EMME, 0 = control) on percentage of ideas from trustworthy pages included in post-test essays with SERP reading times at post-test as a mediator (standardized coefficients).

**Figure 2.** Mediation model for the indirect effect of condition (dummy coded 1 = EMMEs, 0 = control) on document source citations in post-test essays with SERP reading times at post-test as a mediator (standardized coefficients).

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60





1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

## Practitioner Notes

### What is already known about this topic

- Critical reading implies sourcing.
- Sourcing is the process of using information about documents in evaluating and interpreting documents' content.
- Sourcing is related to comprehension of multiple texts on the same topic.
- Interventions have been implemented to improve students' critical reading on the Internet through sourcing.
- Eye movement modeling examples (EMME) can be used to guide students' visual attention while performing a task.

### What this paper adds

- EMME increased students' attention to search engine results page (SERP).
- EMME increased students' attention to author information and texts in trustworthy webpages.
- SERP processing time mediated the effects of EMME on the quality of students' post-test essays.
- More ideas from trustworthy pages and citations of document sources featured students' essays in EMME condition.

### Implications for practice and/or policy

- Videos with experts' gaze replays can be included in tutorials for blended and online learning environments.
- These videos can emphasize the importance of paying attention to specific elements when performing a task or activity.
- These videos can contribute to refine students' metacognitive awareness of the various steps involved in a complex process.