JOURNAL OF
COMMUNICATION

Journal of Communication ISSN 0021-9916

ORIGINAL ARTICLE

Evaluating Scientists as Sources of Science Information: Evidence From Eye Movements

Jessica McKnight & Jason C. Coronel

School of Communication, The Ohio State University, Columbus, OH 43210, USA

In the new information environment, individuals can be exposed to different scientists who disseminate information on scientific topics which may or may not be in the scientist's area of expertise. The current study investigates people's ability to evaluate finer, but critical, distinctions in expertise. We use eye movements and self-report measures to determine the extent to which individuals retrieve, from their memories, professional facts about scientists that signal their area of expertise. Our results suggest that individuals can discern expert from nonexpert scientist sources but self-report measures may not accurately reflect this phenomenon, thus highlighting the value of a converging methods approach. We discuss the theoretical and practical implications of these findings.

Keywords: Science Communication, Source Credibility, Expertise, Eye-Movement Monitoring, Memory.

doi:10.1111/jcom.12317

One of the many challenging tasks confronting individuals in modern societies involves evaluating issues related to science and technology with which they are unfamiliar. In recent years, scientists have become important and prominent communicators of science information to the public, by using outlets such as social media (Dudo & Besley, 2016; Nisbet & Scheufele, 2009; Peters, 2013), and people often turn to scientists as both sources of science information and as aids for comprehending issues related to science (Coleman, 1993; Eiser, Stafford, Henneberry, & Catney, 2009; Feldman, 2013). One consequence of this new information environment is that individuals can be exposed to different scientists, possessing different areas of expertise, who disseminate information on scientific topics which may or may not be congruent with their area of expertise (Oreskes & Conway, 2010).

For example, imagine a person confronts diverging information about vaccines on two different blogposts, each written by a scientist. One blogpost is by a medical doctor and professor of medicine who writes about the efficacy and safety of vaccinations

---

Corresponding author: Jason C. Coronel. e-mail: coronel.4@osu.edu

(Novella, 2016). The author of the second blogpost is a paleontologist and professor of geology who describes the dangers of vaccinations (Scheibner, 2013). As highlighted by this real-life example, this new environment raises important questions about individuals' ability to evaluate finer, but critical, distinctions in expertise. More specifically, when evaluating scientists, do individuals use previously learned information about a scientist's characteristics (e.g., professional qualifications) and recognize that a scientist may not be an expert on the science information he or she is disseminating? Or, do they conclude that scientists are likely to be an expert on the scientific information they choose to disseminate?

The answers to these questions are important given evidence that subject-area experts are more likely to possess greater comprehension of their domain of expertise (Ericsson & Lehmann, 1996; Tetlock, 2005). Thus, if individuals do not differentiate between subject matter experts and nonexperts, then they may have difficulties in determining whether scientific information disseminated by such sources is accurate. The goal of this study, then, is to answer these questions in a unique way by examining the extent to which individuals retrieve from memory information about scientists' professional characteristics (diagnostic of their expertise) while they are evaluating them.

Our study contributes to both the science communication and broader source characteristics literatures in two ways. First, our study examines instances in which individuals evaluate scientists who disseminate science information that is either congruent or incongruent with the scientist's area of expertise. Our work advances the scholarly literature given that previous studies in science communication have only examined instances in which individuals evaluate scientists compared to nonscientists (e.g., scientists vs. a banker or a legal secretary) as sources of science information (Winter & Krämer, 2012; Winter, Krämer, Appel, & Schielke, 2010).

Second, we use for the first time eye-movement monitoring methods in order to determine if individuals retrieve from memory professional facts about scientists (diagnostic of their subject-area expertise) as they disseminate science information at the exact moment the individual is evaluating them. Eye-movement measures possess two unique advantages: (1) They can measure memory retrieval processes at the point in time when these processes are occurring and (2) they can measure these processes without the need for self-report. We build on previous work in cognitive psychology showing that eye movements can be used to infer the kinds of information people retrieve from memory during the retrieval process itself (Altmann, 2004; Richardson & Spivey, 2000). Although we use our novel eye-movement monitoring design in the context of science communication, our paradigm can be used in other domains that investigate the types of information individuals retrieve from memory while evaluating sources (e.g., political or health communication).

Our discussion proceeds as follows. The first section discusses our predictions on how the motivation to distinguish experts from nonexperts can be reflected in memory retrieval of previously learned professional facts about the scientists. We then describe the evidence supporting our claim that eye movements can provide a moment-by-moment measure of memory retrieval processes and their advantages over self-report measures of memory retrieval. We then move to describing our design in which we employ a converging methods approach by using both eye movement and self-report measures of memory retrieval processes. Finally, we show our results and end by discussing the broader implications of our findings for the field of science communication and studies on source credibility.

Memory retrieval and evaluation of scientists

In the current media environment, individuals are able to easily access and learn information about scientists' professional backgrounds either from author information provided at the bottom of any online article or blog post, through social media profiles, or even via a simple Google search.[1]

While evaluating science information disseminated by scientists, individuals may use this previously learned knowledge in order to differentiate them according to their area of expertise. More specifically, people can retrieve previous knowledge about scientists, such as their professional backgrounds, from their memories as scientists disseminate scientific information that may or may not be congruent with their area of expertise. There are at least two theoretical views that make distinct predictions regarding how this memory retrieval process may occur.

One possibility is that individuals, when confronted with information that violates previously learned knowledge, actively monitor the accuracy of their memories by retrieving previously learned facts. Throughout the remainder of this paper, we refer to this perspective as the "strategic monitoring account." Indeed, theoretical and empirical works in cognitive psychology have identified a set of processes ("source monitoring" processes) involved in verifying the validity of one's memories (Johnson, 2006; Johnson, Hashtroudi, & Lindsay, 1993). In addition, certain conditions may increase the likelihood that individuals "check" the accuracy of their memories by engaging in strategic search and retrieval through them (Benjamin, 2011). One such condition involves encountering information that violates previous knowledge or one's expectations as these instances seem to generate greater levels of cognitive processing. For example, a large body of work has shown that information that is inconsistent with one's expectations both elicits greater attention and is better remembered than consistent information (Schmidt, 1985; Stangor & McMillan, 1992). In the context of science communication, this account would predict that individuals will be more likely to retrieve from memory professional facts about scientists when they disseminate science information that is incongruent with their area of expertise (e.g., a paleontologist vs. a medical doctor disseminating information about vaccines).

A second possibility is that the semantic relatedness between the previously learned professional facts about the scientist and the disseminated science information enhances the accessibility of such facts (e.g., a medical doctor disseminating information about vaccines enhances memory retrieval of previously learned information that the scientist earned a Ph.D. in biology). Hereafter, we refer to this theoretical perspective as the "semantic facilitation account." This classic view represents ideas in memory as a network of semantically related concepts or nodes. When one encounters a concept, its corresponding node becomes active and that activation spreads to other surrounding nodes within the network (Collins & Loftus, 1975). A large body of work over the last 30 years has documented the enhancement of accessibility elicited by incoming information on semantically related concepts stored in long-term memory (Meyer & Schvaneveldt, 1971; Neely, 1977; Thompson-Schill, Kurtz, & Gabrieli, 1998). In contrast to the strategic monitoring account, this account predicts that individuals will be more likely to retrieve professional facts about scientists from memory when the scientists disseminate information that is congruent with their area of expertise.

These two accounts make distinct predictions regarding how individuals will retrieve previous knowledge about scientists from their memories as the scientists disseminate scientific information that may or may not be congruent with their area of expertise. We thus have the following two competing hypotheses.

Strategic monitoring

Individuals will retrieve more professional facts about scientists when the scientists disseminate information that is incongruent rather than congruent with their area of expertise.

Semantic facilitation

Individuals will retrieve more professional facts about scientists when the scientists disseminate information that is congruent rather than incongruent with their area of expertise.

Distinguishing between the strategic monitoring and the semantic facilitation accounts is important as these are two very different routes that individuals may take to evaluate scientists who communicate about science topics to the public. In particular, the strategic monitoring account suggests that individuals are actively and intentionally examining the veracity of their memories of a scientist especially when they encounter science information that violates their expectations about the scientist. In contrast, the semantic facilitation account suggests that individuals are passively and (perhaps) automatically accessing prior information about a scientist especially under conditions when there is a semantic match between their prior knowledge of the scientist and the science information he or she is disseminating.

These two accounts comport with prominent theories that specify the manner in which individuals evaluate and digest information from their social environment. In particular, converging evidence across the behavioral sciences suggests that individuals evaluate information either in a heuristic, uncritical, and passive manner or in a deliberative and critical way (Chaiken, 1980; Kahneman, 2011; Lodge & Taber, 2013; Zaller, 1992). These two distinct processes have sometimes been referred to, respectively, as “System 1” and “System 2” (Kahneman, 2011). System 1 processes underlie semantic facilitation (Kahneman, 2011) whereas the processes that support strategic monitoring are components of System 2 processing (Benjamin, 2011). One can argue that in the current information environment, it would be advantageous for individuals to employ a System 2 approach by critically and strategically evaluating information about the content’s source. This is especially important if incorrect or misleading science information is more likely to be communicated by scientists who do not have adequate education or experience on the topic they are disseminating. An individual who intentionally checks their memories of a scientist when they think the scientist may not be an expert on the topic (strategic monitoring) is much more likely to catch potential misinformation compared to an individual who passively accesses their memories when they think a source is an expert on the disseminated science information (semantic facilitation).

Eye-movement measures of memory retrieval

Ideally, communication scholars would like to determine the types of information individuals are retrieving from memory while evaluating a source. But at the critical moment when individuals are evaluating the source there are often no observable physical actions that accompany the retrieval of facts from memory. To study this elusive phenomenon of memory retrieval, researchers have used a variety of self-report measures including recall/recognition measures of memory (Hovland & Weiss, 1951), ratings of perceived source characteristics (Chaiken, 1980; Petty & Cacioppo, 1986), and think-aloud procedures (Williams & Hollan, 1981). These measures, however, have important limitations. Studies usually obtain behavioral measures of memory and ratings of perceived source characteristics after participants have finished evaluating a source and thus may tap into processes that may not have occurred during evaluation (Russo, Johnson, & Stephens, 1989). Think-aloud procedures, which require participants to verbally list what facts they are retrieving from memory, can be used at the time of source evaluation. But asking individuals to explicitly report their memory retrieval processes may change what they retrieve from memory (Wilson & Nisbett, 1978). Thus, determining the types of information individuals retrieve from memory about sources using self-report measures is difficult to do as these processes are occurring and retrospective assessments may introduce bias.

Over the last two decades, a growing area of research in cognitive psychology suggests that memory retrieval processes can be reflected as observable actions via eye movements and as such, researchers can use eye movements as measures of the moment-by-moment processes associated with memory retrieval. One of the main studies used to demonstrate the link between eye movements and memory retrieval is often referred to as the “looking-at-nothing” paradigm. In the classic version of the study, participants watched videos of four talking individuals—each located in one of four quadrants on the screen (Richardson & Spivey, 2000). Each of the heads uttered a fact that participants were instructed to remember (i.e., "Shakespeare's first plays were historical dramas; his last was *The Tempest*"). In a subsequent trial, the researchers removed the heads from the screen. Participants then heard a statement related to one of the previously learned facts ("Shakespeare's first play was *The Tempest*") and were asked to determine whether the statement was true or false. This task required participants to retrieve the previously learned fact from memory. While verifying the statement, the researchers found that participants directed a greater amount of gaze to the location of the previously learned fact (e.g., the location of the head that uttered "Shakespeare's first plays were historical dramas; his last was *The Tempest*") despite the screen being blank (Richardson & Spivey, 2000).

This phenomenon in which individuals, while retrieving information from memory, gaze back at spatial locations previously associated with that information (during encoding) has been observed by other studies across different contexts and domains (Johansson & Johansson, 2014; Platzer, Bröder, & Heck, 2014; Renkewitz & Jahn, 2012; Scholz, Mehlhorn, & Krems, 2016; Scholz, von Helversen, & Rieskamp, 2015; Spivey & Geng, 2001). Collectively, these studies suggest that eye movements can be used as a measure of memory retrieval without the need for self-report responses. Eye movement metrics can thus provide communication scholars with an alternative way to measure memory retrieval processes at the point in time when this retrieval is occurring.

Methods

Participants

We ran a total of 120 participants who were recruited from a large public university in the Midwestern United States and the surrounding community. All participants were compensated with $15 for taking part in the study. We excluded five participants who either fell asleep during our study or were wearing glasses or contacts that interfered with calibration of our eye-tracking instruments. We analyzed data from the remaining 115 participants (60.00% female; Age $M = 24.43$ years, $SD = 7.64$ years, range = 18–74 years).

Materials

Stimuli consisted of four news articles about scientific topics across four unique domains. The scientific topics were thorium fuels, Project Loon, phage therapy, and megascale desalination (see Table 1). We intentionally selected topics which most individuals were unfamiliar with (based on a previous norming task[2]) given that we wanted most of our participants to possess the same level of knowledge and familiarity (i.e., little to none) across all the issues. For each topic, we constructed a news article that described the technology along with the potential benefits and costs associated with its development. The articles were between 346 and 357 words long and each had a headline that was eight words long.

Each of the four scientists were represented by a single color photograph of a White male of middle age, obtained online, that were matched on attractiveness and the

Table 1 Description of Scientific Topics

Thorium Fuels: Nuclear fuels that use the energy contained in the element of Thorium as their basis. Thorium is more abundant than Uranium and produces less harmful radioactive waste.

Phage Therapy: The use of viruses to treat bacterial infections. In some cases, they can potentially be more effective at treating infections than antibiotics.

Project Loon: A development project by tech company Google to spread internet access across the globe by using a series of balloons. The goal is to allow people to access the internet anywhere in the world.

Mega-scale Desalination: The large-scale removal of salts from ocean water. This makes ocean water suitable for drinking and use for farming.

extent to which they were judged to look like a scientist (as determined by a norming task with a different group of participants; attractiveness $M = 4.2$ [1–7 scale]; perceived to look like a scientist $M = 5.11$ [1–7 scale]). We selected males as scientists given that women are underrepresented in fields related to science and engineering (Beede et al., 2011). Each of the scientists was associated with four professional and four nonprofessional facts. There were a total of 16 professional and 16 nonprofessional facts. The professional facts were statements that described career-related qualifications, accomplishments, experiences, etc. (see Table 2). We constructed the professional facts such that four of them (four for each of the four topics) were perceived as characteristics of a scientist who was an expert on one of our four topic areas. For example, a scientist associated with "Has a Ph.D. in Nuclear Engineering" would be perceived as more likely to be an expert on thorium fuels than the other topic areas.

Table 2 Professional Attributes by Topic

Thorium
- Has a Ph.D. in Nuclear Engineering
- Recipient of Top Nuclear Industry Award
- Researches the performance of nuclear energy
- CEO of the Nuclear Energy Institute

Project Loon
- Worked for NASA's scientific balloon team
- Has a Masters in Wireless Communication
- Professor at MIT in Wireless Communications
- Developed high altitude balloons for research

Phage Therapy
- Runs a lab researching antibiotic-resistant infections
- Has a Ph.D. in Molecular Virology
- Winner of the Disease Therapy award
- Has two government-approved viral medications

Mega-scale Desalination
- Published over 100 articles on desalination
- Won Best Paper for desalination research
- Holds patents on water treatment technology
- Has crowd funded a desalination company

The nonprofessional facts referred to noncareer-related activities, hobbies, preferences, etc. (e.g., “Volunteers at a local soup kitchen”). Pretest raters judged eight of the nonprofessional facts as positive characteristics (e.g., “Is a mentor for troubled teens”), the other eight as neutral characteristics (e.g., “Wears a blue tie to work”). All 32 professional and nonprofessional statements were exactly six words long. We included nonprofessional facts for two reasons. First, their inclusion mirrors real-world information environments in which nonprofessional facts tend to also be associated with scientists. Second, they serve as “competitors” to the professional facts. That is, at the time of evaluation, participants are not limited to only one type of information (i.e., professional facts) that they can retrieve from memory about the scientists. We therefore did not intentionally design our study to favor the retrieval of professional facts from memory.

Procedure

We tested participants individually in a quiet room, where they were seated 100 cm away from a 24-inch Asus VG248QE LCD monitor (resolution 1,920 × 1,080), with a refresh rate of 60 Hz. Before the experiment began, we fitted the desktop-mounted SR Research EyeLink 1000 eye tracker and calibrated it for each subject with a 9-point calibration system. We used a chin rest to reduce head movements and a rigid mount to keep the chin and forehead from moving and participants were instructed to remain still throughout each phase of the experiment. We conducted drift correction between each trial. Recordings were taken from the right eye, except for two instances in which reflection off the participant’s glasses necessitated left-eye recording. The entire study took about 1 hour to complete.

Before the start of the experiment, participants took a survey that presented them with information about eight technologies, including the four scientific topics that were used in the main study. We then instructed participants to state the extent to which they supported or opposed the development of each technology and to identify topics with which they were unfamiliar. In order to familiarize participants with all of the science topics, we gave participants brief descriptions of each topic (see Table 1 for the exact descriptions provided to participants).

The experiment consisted of three phases: the learning phase, the sharing phase, and the memory phase. To protect against fatigue, participants had the option to take two 5-minute breaks during the study, one after the learning phase and one after the sharing phase. During the learning phase, we showed participants professional and nonprofessional facts about each of the scientists. Participants were instructed to pay attention to each of the four scientists’ characteristics, as they read news articles shared on social media by these scientists later in the study. Participants were told that they would be viewing and evaluating real articles shared by real scientists, but not that there would be a memory test for the facts about the scientists. A trial consisted of a photo of one of the four scientists shown in the middle of the screen with four empty boxes on both the right and left sides of the screen (see Figure 1). We first showed the photo of the scientist for 6 seconds. Then, we showed the characteristics one at

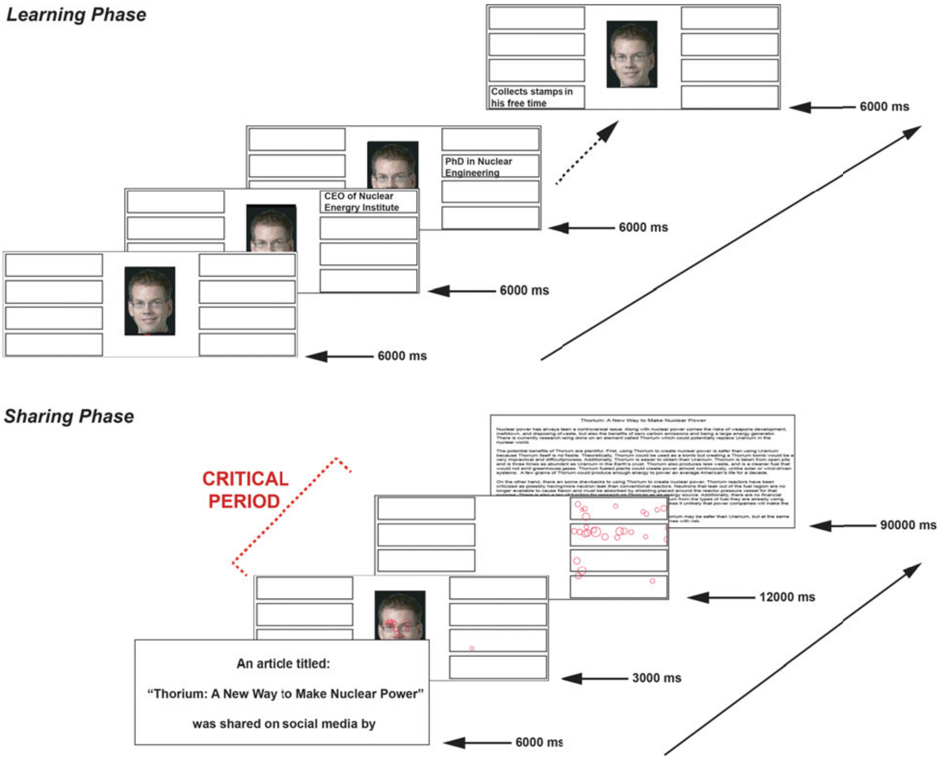

Figure 1 Schematic design of the learning and sharing phases. Eye movements are superimposed in the critical period. Circles represent fixations. Larger circles represent longer fixations. The photo of the scientist here is not the actual photo used in the study.

a time in each of the boxes for 6 seconds (i.e., all four professional facts on one side of the screen and all four nonprofessional facts on the other side of the screen; the location — left or right — of the professional and nonprofessional facts was counterbalanced between participants). After they had learned all the facts about one scientist, participants were then presented facts about the next scientist. Each scientist was associated with professional facts that characterized them as an expert on one of our four scientific domains. The order of presentation of the four trials was randomized across participants.

Next, during the sharing phase, we showed participants four news articles that were purportedly shared on social media by the scientists they learned about in the learning phase. We instructed participants to read each news article. A trial consisted of an article title, shown in the middle of the screen for 6 seconds, that indicated the topic area that would be covered by the article (e.g., "An article titled: 'Phage Therapy: A New Response to Antibiotic Resistance' was shared on social media by") (see Figure 1). Then, participants were shown a photo of the scientist who shared the article for 3 seconds. For two of the scientists, the topic area was congruent with the professional facts that characterized them as an expert on a specific scientific domain. For example, if the article was titled “Phage Therapy: A New Response to Antibiotic Resistance,” then the scientist was one who was previously associated with professional characteristics associated with being an expert on phage therapy.

For the other two scientists, the topic area was incongruent with the professional facts that characterized them as an expert on a specific scientific domain. For example, if the article was titled “Project Loon: Balloon Powered Internet Launches in 2016,” then the scientist was one who was previously associated with professional characteristics on a different domain of expertise (e.g., megascale desalination). The presentation order of the four trials was randomized. The two congruent and incongruent scientists were associated with either positive or neutral nonprofessional facts (e.g., for the two congruent candidates, one was associated with positive facts while the other was associated with neutral facts). We chose positive and neutral nonprofessional facts as we assumed that this characterized many individuals in one’s online social network. The extent to which we assigned an issue or scientist into the congruent or incongruent condition was counterbalanced across conditions.

The photo then disappeared, leaving only the eight empty boxes (four on each side) in which the previous professional and nonprofessional statements were located during the learning phase. This was on the screen for 12 seconds. This 15-second period (starting at the onset of the eight empty boxes and scientist’s photo and stopping when the eight boxes disappeared) was our critical period. If, as we predict, participants are more likely to retrieve professional facts in one condition over another (incongruent > congruent or congruent > incongruent), then based on previous “looking-at-nothing” studies, there should be differences between the conditions in the amount of gaze directed to the four boxes in which the professional facts were located during the learning phase. The region encompassing the four empty boxes that previously contained the professional facts is our critical region of interest.

After this critical period, participants were shown the article for 90 seconds. After the sharing phase, participants answered a survey that asked them to indicate what facts about the scientists they were primarily thinking about in the critical period (i.e., when they saw the scientist’s photo immediately after presentation of the article’s title during the sharing phase). Participants had four response options: the professional facts previously associated with scientists, the nonprofessional facts previously associated with scientists, both professional and nonprofessional facts, or neither professional nor nonprofessional facts. We wanted to examine the extent to which self-report measures of memory retrieval converged or diverged with the outcomes of our eye movement metrics.

Participants then took part in the memory phase. In this phase, we showed participants a fact about the scientist for 4 seconds. Eight of the facts were “old”: the same professional/nonprofessional facts that they saw during the learning phase. Another eight of the facts were “new”: professional/nonprofessional facts that they did not see during the learning phase. The new items were constructed such that they matched the professional and nonprofessional schemas associated with the candidates during the learning phase (e.g., if the scientist was the thorium fuels expert, then the new professional fact was also one that conveyed expertise in thorium). The statement was followed by a photo of the scientist. Once the participants saw the photo, they were instructed to report, via a button press, whether the pairing between the characteristic and the scientist was "old" or "new." Immediately following each old/new decision, participants also made a confidence judgment ("somewhat sure" or "very sure") about their recognition decision. Participants were asked to make their old/new decisions and confidence judgments within a 2-second time window. Presentation order of the trials was randomized.

After the memory phase, participants were given another survey. They were asked to rate the extent to which each scientist was an expert on the topic covered by the article he shared (7 point scale; 1 = Not at all an expert; 7 = Very much an expert). In addition, they answered several demographic-related questions.

Analysis strategy

Of critical interest in this study is the 15-second period (starting at the onset of the eight empty boxes and scientist's photo and stopping when the eight boxes disappeared) during the sharing phase (critical period). If individuals are retrieving previously learned professional facts about the scientists, then we expect gaze to be directed to the four empty boxes that were previously associated with these facts (our critical region of interest). The strategic monitoring and semantic facilitation accounts make distinct predictions regarding the amount of gaze directed to this critical region.

Under the strategic monitoring account, individuals are more likely to check the accuracy of their memories when their expectations are violated. For example, if an article title signaled thorium fuels as the topic domain and the participant expected the thorium expert as the sharer, revealing the phage therapy expert as the sharer will lead them to check if they misremembered the scientists' area of expertise. This requires that they retrieve the professional facts associated with the phage therapy expert. In the congruent case, their expectations match the sharer and there is no immediate rationale for them to reaccess the sharer's professional characteristics. Thus, this account predicts that participants will be more likely to direct their gaze at the critical region in the incongruent than the congruent condition.

According to the semantic facilitation account, the semantic relatedness between the previously learned professional facts about the scientist and the disseminated science information enhances the accessibility of such facts. In our design, this is more likely to occur in the congruent than incongruent condition. In the congruent case, the topic domain of the article and the sharer's domain of expertise are semantically related to each other. In contrast to the strategic monitoring account, then, this account predicts that participants will be more likely to direct their gaze at the critical region in the congruent than the incongruent condition. Thus, by comparing the amount of gaze directed at the critical region for the congruent and incongruent conditions, our design allows us to distinguish between these two competing theoretical accounts.

Results

Eye movements

We created two regions of interest encompassing the two sets of four rectangles (located either on the left or right side of the screen) presented during the sharing phase in which the professional and nonprofessional facts about the scientists were presented during the learning phase. We measured gaze to our critical region of interest (the four rectangles that previously contained the professional facts) as the number of fixations (discrete pauses in the display) directed to that critical region.

If participants monitor the accuracy of their memories (strategic monitoring account) of the scientists by retrieving previously learned professional facts about them, they are more likely to do so in the incongruent case as this is the instance in which their expectations —based on previous memories— are violated. Thus, this account predicts that participants will direct a greater amount of gaze to the critical region in the incongruent than the congruent condition. In contrast, if greater semantic relatedness between the topic domain signaled by the article's title and the sharer's expertise further enhances the accessibility of the professional facts associated with the scientist (via automatic spreading activation)— the semantic facilitation account predicts that participants will direct a greater amount of gaze to the critical region in the congruent than incongruent condition.

To test either of these possibilities, we used congruency as an independent variable (congruent = "1," incongruent = "0") and their fixations to the critical region as a dependent variable (this was a binary variable: "1" indicated that the fixation was located in the professional region while "0" indicated that the fixation was located in the nonprofessional region). We estimated a mixed-effects logistic regression model that treated congruency as a fixed effect and participants and article domain (topic covered by article) as random effects.[3] A significant and positive effect of congruency (see Table 3) suggests that participants were more likely to fixate on the professional (critical) region in the congruent than the incongruent case.[4] This outcome is consistent with the predictions of the semantic facilitation account.

We estimated a second model in which we included a variable that represented the nonprofessional facts about the candidates (positive = "1", neutral = "0"). Although this was counterbalanced across the two conditions within participants, people's perceptions of what is positive could differ. We also included their self-reported ratings of their level of familiarity with the science topic (1 = "I've never heard of this before. This is the first time I've encountered this topic." 5 = "I consider myself an expert on this topic."). This information was obtained before the start of the study. Finally, we also included a 7-item True/False questionnaire (e.g., T/F: Antibiotics will kill viruses as well as bacteria) that has been used in the literature as a measure of general scientific knowledge (National Science Foundation, 2012). Adding these

Table 3 The Effect of Congruency on Gaze to Critical Region

	Model 1	Model 2
Regression estimates		
Congruency	.13*	.12*
	(.06)	(.06)
Positive	—	−.10
		(.06)
Familiarity	—	.18***
		(.05)
Science knowledge	—	.05
		(.06)
Data characteristics		
Number of observations (Fixations)	5,977	5,977
Number of participants	115	115
Number of issues	4	4

Note: Logistic regression coefficients are shown with standard errors in parentheses.
$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

variables did not change our substantive findings (see Table 3). We also found an effect of familiarity in which participants were more likely to direct their gaze to the critical region for topics in which they had a high degree of familiarity.[5]

Recognition memory

We analyzed the behavioral data from the recognition memory phase of the study to rule out an alternative explanation of the eye-movement results: Participants had better memories of the professional facts for the congruent than incongruent scientists and, thus, were more likely to engage in memory retrieval. In order to assess recognition memory, we calculated the hit rate (proportion of old items correctly classified as "old") and the false-alarm rate (proportion of new items incorrectly classified as "old") for the professional facts associated with the scientists. Both congruent (hit rate $M = 77\%$, $SD = 14\%$; false-alarm rate $M = 30\%$, $SD = 19\%$) and incongruent (hit rate $M = 74\%$, $SD = 17\%$; false-alarm rate $M = 28\%$, $SD = 18\%$) scientists elicited similar hit and false-alarm rates. We assessed recognition memory using the discriminability index d-prime, which takes into account information about hit and false-alarm rates (Macmillan & Creelman, 2005). A d-prime score above 0 suggests that participants can reliably discriminate between old and new items. A score of 0 suggests that participants are unable to discriminate between old and new items. Participants were able to reliably discriminate between old and new items for both congruent ($M = 1.44$, $SD = .69$; $t(114) = 22.57$, $p < .001$) and incongruent ($M = 1.42$, $SD = .78$; $t(114) = 19.56$, $p < .001$) scientists and there were no differences in the d-prime scores between the two within-subject conditions ($t(114) = .29$, $p = .77$). These results suggest that participants had intact memories for the professional facts they learned and that neither the congruent nor incongruent scientists were easier to encode and retrieve than the other.

Self-reported measures

Finally, we conducted exploratory analyses in order to examine the extent to which eye-movement results can be dissociated from self-report responses. We measured participants' self-reported responses in which they indicated what facts about the scientists they were primarily thinking about when they saw the scientist's photo immediately after presentation of the article's title during the sharing phase. Participants were given four options: the professional facts previously associated with scientists, the nonprofessional facts previously associated with scientists, both professional and nonprofessional facts, and neither professional or nonprofessional facts. This analysis allowed us to examine the extent to which participants' introspective assessments of what facts about the scientists they were accessing converged or diverged with the results of the eye movement metrics. On average, participants self-reported that they were predominantly thinking about the scientists' professional facts (professional = 66% of responses, nonprofessional = 13%, both = 12%, neither = 9%). Critically, there were no differences in the rate at which participants reported thinking about the professional facts between the congruent ($M = 67\%$) and incongruent scientists ($M = 65\%$, $t(114) = .45$, $p = .65$). We also estimated a multinomial logit model (with robust standard errors clustered on participants) in order to retain all four categories with "professional facts" as the contrast category. Coefficient estimates for the full multinomial logit are displayed in Table 4. Here, we again neither see significant effect of congruency nor an effect of familiarity.

Next, we examined the extent to which our classification of congruent and incongruent scientists predicted participants' self-reported assessments about the scientists' level of expertise. The self-reported subjective assessments were obtained after the memory phase and asked participants to rate the extent to which a scientist was an expert on the topic covered by the article he shared (7-point scale; 1 = Not at all an expert; 7 = Very much an expert). We used congruency as an independent variable and their self-reported ratings of a scientist's level of expertise as a dependent variable. We estimated a mixed-effects logistic regression model with congruency (congruent = "1", incongruent = "0") as a fixed effect and participants and article domain (topic covered by article) as random effects. A significant and positive effect of congruency ($\beta = .35$, $SE = .09$, $p < .001$) suggests that participants were able to correctly identify the experts from the nonexperts given the articles they shared. However, these data cannot directly speak to the question of whether they made this distinction during the moment they were evaluating the scientists (i.e., as the scientists shared the article) as these measures were obtained after the point of evaluation.

Discussion

In the new information environment, individuals can be exposed to different scientists, possessing different areas of expertise, who disseminate information on scientific

Table 4 The Effect of Congruency on Self-Reported Responses of Facts Retrieved from Memory

	Nonprofessional	Both	Neither
Model 1			
Regression estimates			
Congruency	−.29	−.03	.16
	(.24)	(.24)	(.25)
Data characteristics			
Number of observations = 460			
Model 2			
Regression estimates			
Congruency	−.30	−.03	.18
	(.24)	(.25)	(.25)
Positive	.46	.05	−.24
	(.28)	(.26)	(.25)
Familiarity	.13	.03	.26
	(.21)	(.25)	(.18)
Science knowledge	−.18	−.16	.17
	(.13)	(.12)	(.23)
Data characteristics			
Number of observations = 460			

Note: Cell entries are multinomial logistic regression coefficients with “professional facts” as the contrast category. Robust clustered standard errors are in parentheses.
$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

topics which may or may not be congruent with their area of expertise. Previous studies have largely focused on instances in which individuals evaluate scientists compared to nonscientists as sources of science information (Winter & Krämer, 2012; Winter et al., 2010). In these instances, it is relatively easy for individuals to distinguish experts from nonexperts (e.g., scientists vs. bankers). We investigated the extent to which individuals could make fine, but critical, distinctions between scientists' domain of expertise. We used eye movements to determine the extent to which people retrieve, from their memories, professional facts about scientists that signal their level of expertise. We also employed self-reported measures of facts retrieved from memory. We reasoned that individuals either actively and intentionally examine the veracity of their memories of a scientist (strategic monitoring account) or passively and (perhaps) automatically access prior information about a scientist (semantic facilitation account).

We found that participants were more likely to direct their gaze to the critical region (area where professional facts were located during the learning phase) in the congruent than incongruent condition. This outcome is consistent with the predictions of the semantic facilitation account. That is, the semantic relationship between the scientific domain invoked by the article's title and scientists' domain of expertise was stronger in the congruent than incongruent condition—thus facilitating enhanced accessibility and retrieval of professional facts via automatic spreading activation. This result is likely not due to encoding differences of the professional facts between the two conditions as participants demonstrated equally robust recognition memories for congruent and incongruent scientists. Although these results indicate that individuals were assessing the scientist's level of expertise and using their previous knowledge about them at the moment of evaluation, there are reasons to believe that individuals who comport with the predictions of the strategic monitoring account are preferable to ones who behave in accordance to the predictions of the semantic facilitation account. Memory is highly vulnerable to errors (for reviews, see Gallo, 2013; Wells & Loftus, 2003) and it is advantageous for people to assess the degree to which they possess accurate memories of sources of science information. In the new information environment, scientists who do not have adequate education on the topic they are disseminating may be more likely to disseminate misinformation. An individual who examines the veracity of their memories when they believe a scientist may not be an expert on the topic is much more likely to catch such misinformation.

Critically, we also found a dissociation between self-report and eye-movement measures. In particular, the self-report measures indicated that on average, participants were equally likely to access and think about the professional characteristics between the congruent and incongruent scientists—an outcome that does not match the results from the eye-movement measures. One explanation for this result is that this self-reported information was not obtained at the moment when individuals were evaluating the scientists. Thus, their retrospective evaluations may have led to an inaccurate assessment.

Furthermore, our study also highlights the value of a converging methods approach. That is, using different methods to measure the same concept. Our other self-report measures suggest that participants were able to make accurate assessments about the scientists' level of expertise over a given domain. One concern is that these data were obtained after the point of evaluating the scientist and there is the possibility that they did not employ this information during the moment they were evaluating the scientists (i.e., as the scientists shared the article). However, additional eye-movement data indicating that participants were distinguishing experts from nonexperts via differences in the access of professional facts provide supporting evidence to the inferences drawn from the self-report measures.

One advantage of our design is that it incorporated features that most social media channels share. For example, the materials shown to participants in this study represent the types of information about scientists that are readily available to individuals through social media sites or the Internet in general. Participants viewed a photo or profile picture of the scientist, information about their professional background, and information about their personal characteristics. We also showed participants an article that each scientist chose to share with their social network and the public. These are the same types of data that people could use on social media in order to evaluate a scientist as a source of science information. Our design provides an approximation of how these evaluative processes may occur without showing the relevant information in the format of a specific social media or news media environment. Because the design of this experiment does not visually look like a particular website, the processes observed could thus occur in any online environment where these types of information are available.

Although this study is a useful first step, caution is warranted in terms of generalizing some of the study's findings. For example, the materials used in this study were constructed to limit confounding factors, such as familiarity and topic salience. Consequently, the topic domains are not representative of the broad array of scientific topics that populate the current information environment. Indeed, our exploratory analyses revealed that familiarity with the science information makes it more likely for individuals to retrieve professional facts about scientists. Future work should investigate the extent to which the effects we observe here are also obtained for highly salient and affectively charged scientific topics. Furthermore, our participants are not a representative sample as all of them have received or are currently receiving a college education. Thus, they may be on the higher end of the spectrum in terms of knowledge and interest in science-related information.

Future work should also examine information literacy interventions as a way of assisting individuals in critically and strategically evaluating scientists who are sharing scientific information online. One promising intervention is the checklist approach, wherein checklists are developed to guide users through the credibility evaluation process and individuals are taught to ask and answer a list of questions about information they encounter online (e.g., Kurtz-Rossi & Duguay, 2010). Given the findings from this study, some potential items that could be included on the checklist for evaluating science sources online are: (1) Who is the person who shared the content? (2) What is that person's professional background? (3) Does that person's professional background match the content topic?

Finally, our study further highlights the utility of eye-movement measures in investigations of information sources. As shown here, communication scholars interested in determining the types of information individuals retrieve from memory at the exact moment that they are evaluating a source will find eye-movements useful. Eye-movement measures can measure moment-by-moment retrieval processes and can do so without the need for self-report.

In summary, our results suggest that people can distinguish expert from nonexpert sources of science communications at the moment of evaluation even when both sources are scientists. The semantic relatedness between the previously learned professional facts about the scientist and the disseminated science information enhances the accessibility of such facts diagnostic of expertise. However, the fact that individuals in our study do not seem to check the veracity of their memory of the scientists when their expectations are violated has potential implications in more complex environments in which scientists disseminate misinformation on topics in which they are not experts. This is an important area for future research to explore; as scientists are increasingly encouraged to communicate directly with the public, more scientists are likely to communicate about science topics in which they have low levels of expertise.

Acknowledgments

The authors are grateful to Michael Sheehan and Jared MacDonald for their help during data collection and their work on creating the stimuli used in this study. They also thank Molly Ritchie, Matthew Sweitzer, and graduate students enrolled in Communication Research Methods for their helpful feedback and comments. The authors received no financial support for the research, authorship, and/or publication of this article. The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Notes

1 Throughout the discussion in this paper, we assume that individuals have already learned information about a scientist's professional background by the time they are evaluating the scientific information they are disseminating.

2 By "norming task" we are referring to a task in which a different group of participants — recruited from Amazon's Mechanical Turk — rated a larger set of our various stimuli (issues, faces of the scientists, professional facts). For example, in the case of the scientists' faces, respondents rated 20 faces on dimensions such as attractiveness and the extent to which they looked like a scientist. We then selected four faces (that were rated equally in those dimensions) and those became the stimuli we used for the study.

3 For more information on mixed-effects models, see Baayen, Davidson, and Bates (2008).

4 The odds ratio for the congruency coefficient is 1.14. This suggests that participants are 1.14 times more likely to direct their gaze to the region containing the professional facts in the congruent than incongruent condition.

5 The same pattern of results was obtained when gender, age, and SES were included in the model. To measure SES, we used the highest level of education completed by the parents.

References

Altmann, G. T. (2004). Language-mediated eye movements in the absence of a visual world: The "blank screen paradigm". *Cognition*, **93**(2), B79–B87.

Baayen, R. H., Davidson, D., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, **59**(1), 390–412 https://doi.org/10.1016/j.jml.2007.12.005

Beede, D. N., Julian, T. A., Langdon, D., McKittrick, G., Khan, B., & Doms, M. E. (2011). Women in STEM: A gender gap to innovation. *Economics and Statistics Administration*, 04–11.

Beede, D.N., Julian, T.A., Langdon, D., McKittrick, G., Kahn, B., & Doms, M.E. (2011). *Women in STEM: A gender gap to innovation* (Issue Brief No. 04-11). Washington, DC: U.S. Department of Commerce.

Chaiken, S. (1980). Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of Personality and Social Psychology*, **39**(5), 752.

Coleman, C.-L. (1993). The influence of mass media and interpersonal communication on societal and personal risk judgments. *Communication Research*, **20**(4), 611–628.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, **82**(6), 407.

Dudo, A., & Besley, J. C. (2016). Scientists' prioritization of communication objectives for public engagement. *PloS One*, **11**(2), e0148867.

Eiser, J. R., Stafford, T., Henneberry, J., & Catney, P. (2009). "Trust me, I'm a scientist (not a developer)": Perceived expertise and motives as predictors of trust in assessment of risk from contaminated land. *Risk Analysis*, **29**(2), 288–297.

Ericsson, K. A., & Lehmann, A. C. (1996). Expert and exceptional performance: Evidence of maximal adaptation to task constraints. *Annual Review of Psychology*, **47**(1), 273–305.

Feldman, L. (2013). Breaking boundaries|cloudy with a chance of heat balls: The portrayal of global warming on the daily show and the Colbert report. *International Journal of Communication*, **7**, 22.

Gallo, D. (2013). *Associative illusions of memory: False memory research in DRM and related tasks*. New York: Psychology Press.

Hovland, C. I., & Weiss, W. (1951). The influence of source credibility on communication effectiveness. *Public Opinion Quarterly*, **15**(4), 635–650.

Johansson, R., & Johansson, M. (2014). Look here, eye movements play a functional role in memory retrieval. *Psychological Science*, **25**(1), 236–242.

Johnson, M. K. (2006). Memory and reality. *American Psychologist*, **61**(8), 760.

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin*, **114**(1), 3.

Kahneman, D. (2011). *Thinking, fast and slow*. New York: Macmillan Publishers.

Kurtz-Rossi, S., & Duguay, P. (2010). Health information literacy outreach: Improving health literacy and access to reliable health information in rural Oxford County, Maine. *Journal of Consumer Health on the Internet*, **14**(4), 325–340.

Lodge, M., & Taber, C. S. (2013). *The rationalizing voter*. Cambridge, England: Cambridge University Press.

Macmillan, N. A., & Creelman, C. D. (2005). *Detection theory: A user's guide*. (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology*, **90**(2), 227.

National Science Foundation. (2012). *Science and technology: Public attitudes and understanding* (Science and Engineering Indicators 2012). Washington, D.C. Retrieved from https://www.nsf.gov/statistics/seind12/pdf/c07.pdf

Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General*, **106**(3), 226.

Nisbet, M. C., & Scheufele, D. A. (2009). What's next for science communication? Promising directions and lingering distractions. *American Journal of Botany*, **96**(10), 1767–1778.

Novella, S. (2016, 2 November). Measles more deadly than previously thought. Retrieved from https://sciencebasedmedicine.org

Oreskes, N., & Conway, E. M. (2010). *Merchants of doubt: How a handful of scientists obscured the truth on issues from tobacco smoke to global warming*. New York: Bloomsbury Publishing Inc.

Peters, H. P. (2013). Gap between science and media revisited: Scientists as public communicators. *Proceedings of the National Academy of Sciences*, **110**(Supplement 3), 14102–14109.

Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. In *Communication and persuasion* (pp. 1–24). New York: Springer.

Platzer, C., Bröder, A., & Heck, D. W. (2014). Deciding with the eye: How the visually manipulated accessibility of information in memory influences decision behavior. *Memory & Cognition*, **42**(4), 595–608.

Renkewitz, F., & Jahn, G. (2012). Memory indexing: A novel method for tracing memory processes in complex cognitive tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **38**(6), 1622.

Richardson, D. C., & Spivey, M. J. (2000). Representation, space and Hollywood squares: Looking at things that aren't there anymore. *Cognition*, **76**(3), 269–295.

Russo, J. E., Johnson, E. J., & Stephens, D. L. (1989). The validity of verbal protocols. *Memory & Cognition*, **17**(6), 759–769.

Scheibner, V. (2013, 18 January). Measles vaccines part I: Ineffectiveness of vaccination and unintended consequences. Retrieved from http://www.vaccinationcouncil.org

Schmidt, S. R. (1985). Encoding and retrieval processes in the memory for conceptually distinctive events. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **11**(3), 565.

Scholz, A., Mehlhorn, K., & Krems, J. F. (2016). Listen up, eye movements play a role in verbal memory retrieval. *Psychological Research*, **80**(1), 149–158.

Scholz, A., von Helversen, B., & Rieskamp, J. (2015). Eye movements reveal memory processes during similarity- and rule-based decision making. *Cognition*, **136**, 228–246.

Spivey, M. J., & Geng, J. J. (2001). Oculomotor mechanisms activated by imagery and memory: Eye movements to absent objects. *Psychological Research*, **65**(4), 235–241.

Stangor, C., & McMillan, D. (1992). Memory for expectancy-congruent and expectancy-incongruent information: A review of the social and social developmental literatures. *Psychological Bulletin*, **111**(1), 42.

Tetlock, P. (2005). *Expert political judgment: How good is it? How can we know?* Princeton, NJ: Princeton University Press.

Thompson-Schill, S. L., Kurtz, K. J., & Gabrieli, J. D. (1998). Effects of semantic and associative relatedness on automatic priming. *Journal of Memory and Language*, **38**(4), 440–458.

Wells, G. L., & Loftus, E. F. (2003). Eyewitness memory for people and events. In A. M. Goldstein (Ed.), *Handbook of psychology: Forensic psychology*. (pp. 149–160). New York: Wiley.

Williams, M. D., & Hollan, J. D. (1981). The process of retrieval from very long-term memory. *Cognitive Science*, **5**(2), 87–119.

Winter, S., & Krämer, N. C. (2012). Selecting science information in Web 2.0: How source cues, message sidedness, and need for cognition influence users' exposure to blog posts. *Journal of Computer-Mediated Communication*, **18**(1), 80–96.

Winter, S., Krämer, N. C., Appel, J., & Schielke, K. (2010). Information selection in the blogosphere: The effect of expertise, community rating, and age (pp. 802–807). In

Proceedings of the 32nd Annual Conference of the Cognitive Science Society. Austin, TX: Cognitive Science Society.

Wilson, T. D. & Nisbett, R. E. (1978). The accuracy of verbal reports about the effects of stimuli on evaluations and behavior. *Social Psychology*, **41**(2), 118–131.

Zaller, J. R. (1992). *The nature and origins of mass opinion*. Cambridge, England: Cambridge University Press.