**ABSTRACT**

**Aim/Purpose**
To examine how positive/negative message framing – based on peripheral cues (regarding popularity, source, visuals, and hyperlink) – affects perceptions of credibility of scientific information posted on social networking sites (in this case, Facebook), while exploring the mechanisms of viewing the different components.

**Background**
Credibility assessment of information is a key skill in today's information society. However, it is a demanding cognitive task, which is impossible to perform for every piece of online information. Additionally, message framing – that is, the context and approach used to construct information—may impact perceptions of credibility. In practice, people rely on various cues and cognitive heuristics to determine whether they think a piece of content is true or not. In social networking sites, content is usually enriched by additional information (e.g., popularity), which may impact the users' perceived credibility of the content.

**Methodology**
A quantitative controlled experiment was designed (N=19 undergraduate students), collecting fine grained data with an eye tracking camera, while analyzing it using transition graphs.

**Contribution**
The findings on the mechanisms of that process, enabled by the use of eye tracking data, point to the different roles of specific peripheral cues, when the message is overall peripherally positive or negative. It also contributes to the theoretical literature on framing effects in science communication, as it highlights the peripheral cues that make a strong frame.
Eye-Tracking Study of Credibility of Scientific Information

Findings The positively framed status was perceived, as expected from the Elaboration Likelihood Model, more credible than the negatively framed status, demonstrating the effects of the visual framing. Differences in participants' mechanisms of assessing credibility between the two scenarios were evident in the specific ways the participants examined the various status components.

Recommendation for Practitioners As part of digital literacy education, major focus should be given to the role of peripheral cues on credibility assessment in social networking sites. Educators should emphasize the mechanisms by which these cues interact with message framing, so Internet users would be encouraged to reflect upon their own credibility assessment skills, and eventually improve them.

Recommendation for Researchers The use of eye tracking data may help in collecting and analyzing fine grained data on credibility assessment processes, and on Internet behavior at large. The data shown here may shed new light on previously studied phenomena, enabling a more nuanced understanding of them.

Impact on Society In an era when Internet users are flooded with information that can be created by virtually anyone, credibility assessment skills have become ever more important, hence the prominence of this skill. Improving citizens' assessment of information credibility — to which we believe this study contributes — results on a greater impact on society.

Future Research The role of peripheral cues and of message framing should be studied in other contexts (not just scientific news) and in other platforms. Additional peripheral cues not tested here should be also taken into consideration (e.g., connections between the information consumer and the information sharer, or the type of the leading image).

Keywords credibility assessment, message framing, social networking sites, peripheral cues, eye tracking

INTRODUCTION

Credibility consumers’ connection to a given message — or the extent to which they believe it — is key to its comprehension and acceptance (Heesacker, Petty, & Cacioppo, 1983). Indeed, critical thinking, which credibility assessment is an integral part of, is considered to be one of the most prominent skills in today's information age (Levy & Ramim, 2016). However, assessing credibility is a demanding cognitive task, and as people are constantly exposed to huge amounts of information streams online in a plethora of digital channels, it is unreasonable to expect them to implement a thorough credibility assessment process for every piece of information they see. Rather, people rely on various cues and cognitive heuristics to determine whether they think a piece of content is true or not, especially when they are not familiar with the subject matter (Fogg, 2003; Metzger, 2007; Petty & Cacioppo, 1986; Sundar, 2008).

Social Networking Sites (SNS) — like Facebook, Instagram, or YouTube — are amongst the most popular websites for young Internet users (Smith & Anderson, 2018). As such, they may serve as catalysts for learning, with users interacting with streams of knowledge in new ways (Buchem, 2011; Haugsbakken & Langseth, 2014; Kop, 2012). These platforms are characterized by a very large volume of activity, which translates into a very large volume of shared and consumed content.

Originally intended for social interactions, SNS have developed into major channels of news consumption (Gottfried & Shearer, 2017; Hermida, Fletcher, Korell, & Logan, 2012; Kwak, Lee, Park, & Moon, 2010; Tandoc & Johnson, 2016). However, as opposed to traditional news venues (e.g., newspapers or television), SNS users are often actively sharing news, adding their own (often subjective) take on the original content, which makes credibility assessment extremely challenging (Bucchi, 2017;
Sharon, Ryder, Osborne, Laslo, & Swirski, 2017). That is, SNS often present their users with a blend of mass and interpersonal communication, with no clear boundaries between that two — a result of sharing (and re-sharing) the original content and adding additional layers of information on top of it (Kröl & Wiśniewska, 2017; Neubaum & Krämmer, 2017).

Content posted on SNS is usually enriched by additional information, automatically added by the website, such as details about the user who posted it (most commonly including username and profile image), measures of popularity, and/or indications of publication time. These peripheral cues might impact users' perceived credibility of the content (Hayat & Hershkovitz, 2018; Hayat, Hershkovitz, & Azaran, 2019; Hershkovitz & Badarneh, 2018). This is explained by the Elaboration Likelihood Model (ELM) (Petty & Cacioppo, 1986), according to which people take a peripheral route when evaluating content they are not familiar with or are not motivated to evaluate. For example, content popularity — easily visible on Facebook by the number of Likes, Shares, or Comments — may elicit "bandwagon heuristics", according to which "if others think that this is a good story, then I should think so too" (Sundar, 2008, p. 83). This theory is backed up by recent studies demonstrating the effects of peripheral cues on credibility assessment in the social media (Granjon & Benedic, 2017; Hayat & Hershkovitz, 2018; Hayat, Hershkovitz, & Azran, 2019; Hershkovitz & Badarneh, 2018; Huang, 2015; Lin, 2016; Waddell, 2018; Zhou, 2012).

In addition, message framing may have an effect on how people consume the information embodied in the message. Framing is the context and approach used to construct information. Different framings have been repeatedly found to have an effect on message acceptance (Gallagher & Updegraff, 2012; Schütz & Wiedemann, 2008). In the context of SNS, framing may also be associated with the peripheral cues surrounding a message; for example, posting same text messages with different indications of popularity (e.g., high/low) may impact message acceptance (Hershkovitz & Badarneh, 2018). However, there is still lack of research into the mechanisms involved in this process. The current research aims at bridging this gap.

The purpose of this study was to examine how message framing, based on peripheral cues, affects credibility; whereas an updated literature review of this understudied association is depicted in the next sub section 1.1. The use of an eye tracking camera allowed to explore this phenomenon at the granularity of specific peripheral cues. Sub section 1.2 reviews the use of this methodology within the context of credibility assessment.

**Message Framing and Credibility Assessment**

It has been repeatedly demonstrated that different framings impact perceptions of credibility (e.g., Cobb, 2005; Detenber, Ho, Ong, & Lin, 2018; O’Keefe & Jensen, 2008; Roh, Rickard, McComas, & Decker, 2018; Schütz & Wiedemann, 2008; Smith & Petty, 1996). Of special significance are "positive" and "negative" framings (Chong & Druckman, 2007).

Researchers find bidirectional associations between positive emotions and message popularity. On the one hand, the appearance of positive emotions in a post leads to this post being more popular; on the other hand, people feel more positive emotions as their posts become popular (Bazarova, Choi, Sosik, Cosley, & Whitlock, 2015; Kite, Foley, Grunseit, & Freeman, 2016). It has also been established that design is strongly associated with emotions and attitudes, where appealing (or unappealing) designs eliciting positive (or negative) emotions (Jang & Namkung, 2009; Lee, Ha, & Widdows, 2011; Norman, 2003; Simon, Brexendorf, & Fassnacht, 2013). Emotions expressed in posts to which SNS users are exposed may even be contagious, increasing the positive or negative feelings of the user (Kramer, Guillory, & Hancock, 2014), and this, in turn, may affect information processing (Nabi, 1999; Tiedens & Linton, 2001). In short, various peripheral cues of a Facebook post (e.g., number of Likes and Shares), may increase credibility assessments of that post by increasing positive emotions towards it (Maathuis, Rodenburg, & Sikkel, 2004; Mileti, Prete, & Guido, 2013; Nabi, Gustafson, & Jensen, 2018; Söderlund & Rosengren, 2007; Wang & Hickerson, 2016).
**Studying Credibility Assessment Using Eye tracking**

Traditional credibility studies mostly rely on assessment tasks and self-reporting methods. But Internet studies, specifically in the fields of human computer interaction and usability, opt for eye tracking data to explore actual behavioral patterns (Poole & Ball, 2005). Therefore, it is not surprising that the latter methodology has emerged in studies of online credibility and trust (e.g., Kammerer & Gerjets, 2012; Lee & Pang, 2017).

Tracking a person’s eye movement while conducting a credibility assessment task, makes it possible to explore two aggregated measures of the evaluation process. First, at what the person was looking can be easily associated with specific cues on the screen. Second, how long the person was engaged with each cue (and with the whole task) is immediately obvious. These two components are directly associated with theories of credibility assessment, specifically with the Prominence-Interpretation theory (Fogg, 2003). Based on this theory, credibility assessment depends on two processes: noticing something and interpreting it. The analysis of location and duration enabled by information obtained from eye tracking allows to understand what was prominent and to estimate interpretation efforts, as longer gazes are indicative of a higher level cognitive processing (Follet, Le Meur, & Baccino, 2011).

For example, in a study of credibility evaluation of online information, the authors measured, among other variables, time spent on reading the references section of a blog post; this measure was used as an indicator of interest in evaluating information based on the possibilities afforded by this section (Glasheen, 2013).

In addition, based on locations of gazes and their timestamps, it was possible to induce the path a person took while engaged in an evaluation process. This too may be directly associated with the Prominence-Interpretation theory (Fogg, 2003) - as based on this theory, the evaluation process is multi step. Deciphering the series of elements examined by a person illuminates the multiple steps he or she took, and when they stopped the process and made a decision. This path can then be analyzed and visualized in various ways, benefitting different kinds of explorations (Eraslan, Yesilada, & Harper, 2015; Peysakhovich & Hurter, 2018). In the past, researchers have interpreted behavioral patterns in Web searches by looking at the use of similar paths without using eye tracking (e.g., Cartright, White, & Horvitz, 2011; Xie & Joo, 2010). To the best of our knowledge, the process of credibility assessment has not yet been considered by examining empirical data on the path a person takes while conducting credibility assessment. To gain a more nuanced understanding of the evaluation process, this methodological approach was selected.

**Research Questions**

To meet the goals of this study, the following research questions on the credibility assessment of scientific status on Facebook were formulated:

1. Which components of the status are reported by participants as important for credibility assessment?
2. Which components of the status are viewed in practice during a credibility assessment task, when statuses are either positively or negatively framed using only peripheral cues?

**Methodology**

**Research Population**

Participants were 19 undergraduate students from a large public university in the center of Israel. Participants were students of various disciplines, including Exact Sciences, Life Sciences, Social Sciences, and the Humanities; they were recruited on campus while studying in the library (all were recruited during a single day) – the authors contacted students arbitrary at the library, with the inclusion criteria being B.A. students and knowledgeable with the language in which the information was pre-
Research Design

To study the ways people assess differently framed scientific content on Facebook, where framing is based on peripheral cues, a controlled experiment was designed, in which each participant was exposed to two statuses, one of which was framed positively and the other negatively. Hence, this is a within-subject design, with the manipulated independent variable being the post framing (positive/negative); whereas this approach was taken, due to the fact that different people may present different credibility assessment behaviors (derived from different cognitive, meta cognitive and affective processes involved in the task).

The participants were shown two mock statuses, made up by the authors and presented as if they were posted on Facebook. It is important to note that as this study was about perceived credibility, the question of whether the content presented in the statuses was true or false was irrelevant.

The "positive" status presented information on wine produced from vines grown from the seed of an ancient Assyrian vine found in archeological diggings. The "negative" status presented information on the extraction of a component in moles’ defecation that reduces their pain when they are injured and mentioned that this component might be used to develop effective painkillers for humans.

The language of both statuses had a similar level of difficulty, and the postings had similar lengths; both were presented in a popular science style, that is, in a way accessible to the layperson, with no jargon; both were presumably posted about the same time (a two day difference). Lastly, both presented topics to which it was assumed that laypersons could connect, and of which they would have a similar knowledge (or ignorance) level. This assumption of "similarity" was supported by the collected data (see the Findings section).

Figure 1: Positively framed (left) and negatively framed (right) statuses used here

As mentioned, the positive and negative framing was based on peripheral cues; we used components normally presented in Facebook alongside any status to determine differences:

- Number of Likes, Comments, and Shares: higher for the positive status;
Eye-Tracking Study of Credibility of Scientific Information

- Username and profile image of the person who posted the status: appealing image and scientifically associated name for the positive status, Facebook default image and general name for the negative status;
- Leading image: appealing for the positive status, unappealing for the negative status (for matter of simplicity, we will refer to it as "image");
- Nature of hyperlinks added: science.org vs. health4u.co.il for the positive vs. negative statuses, respectively.

Images of the statuses are presented in Figure 1.

Detailed differences between the two posts are presented in Table 1. Indeed, the positively framed messages were perceived as more credible as the negatively framed message (recall that the text was identical in both cases) (see the Findings section).

Table 1: Characteristics of the two statuses used in the study

<table>
<thead>
<tr>
<th>Component</th>
<th>Positively framed Status: Assyrian Wine</th>
<th>Negatively framed Status: Mole Defecation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Text</td>
<td><em><strong>Drinking wine like in the Assyrian period</strong></em> Grape seeds from the Assyrian period, discovered about a decade ago in archeological diggings in the region of Florence (Italy), were successfully sprouted by a research team of Siena University. Seven years later, the Assyrian vines were already giving fruits, in the form of red grape clusters. The process of producing wine from these grapes was successfully completed a few weeks ago, after it has been aged in wood barrels for a few months. Now, people from both the scientific community and the wine community are waiting to taste the Assyrian wine. <a href="https://www.science.org/wine.html">https://www.science.org/wine.html</a></td>
<td>Soon to come: ParacetaMole? Researchers from Georgetown University (USA) were able to isolate a painkilling material from moles' defecations. More than twenty years ago, biochemists in different parts of the world had noticed that moles use their defecations to help their offspring after being wounded. Now, for the first time, the researchers were able to isolate the component in charge of reducing the pain. The head of the research team said that this component is considered to be combined in painkillers, as it is cheap and highly available. <a href="https://www.health4u.co.il/acamole.html">https://www.health4u.co.il/acamole.html</a></td>
</tr>
<tr>
<td>User Profile Image</td>
<td>Shiny objects pop up from a person's brain</td>
<td>Facebook Default</td>
</tr>
<tr>
<td>Username</td>
<td>Science for All</td>
<td>Feel Well</td>
</tr>
<tr>
<td>Attached Hyperlink Do-</td>
<td>science.org</td>
<td>health4u.co.il</td>
</tr>
<tr>
<td>main</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leading Image</td>
<td>Shiny, attractive red grapes, held between a person's healthy looking palms</td>
<td>A weirdly appearing mole, on top of a Paracetamol package</td>
</tr>
<tr>
<td># Likes</td>
<td>42,518</td>
<td>1</td>
</tr>
<tr>
<td># Comments</td>
<td>192</td>
<td>1</td>
</tr>
<tr>
<td># Shares</td>
<td>73</td>
<td>-</td>
</tr>
</tbody>
</table>
Both an eye tracking camera and a research questionnaire (pen and paper) were used to collect data from the participants. The eye tracking camera was used to capture participants’ gaze at the screen while they were conducting a credibility assessment task; the questionnaire was used to collect participants’ self-reported data on that task and on the other research variables.

**Eye tracking camera**

Eye Tribe Tracker™ was used, which is a low cost portable eye tracking camera manufactured by the Eye Tribe™ company, whereas data from the eye tracking sessions was collected using Eyeproof™, an online software developed by the same company. Although the device has a relatively low sampling frequency level (60 Hz), its accuracy — within the context of Social Science research — is comparable to more sophisticated, higher frequency devices (Dalmaijer, 2014; Funke et al., 2016; Ooms, Dupont, Lapon, & Popelka, 2015). Note that a frequency of 60 Hz means that every second, 60 data points are collected.

**Research questionnaire**

The pen and paper questionnaire consisted of three parts. In the first part, participants were asked to report on certain demographic variables (age, gender, and the faculty where they studied) and on their Facebook use characteristics (extent of use, active/not active, number of Facebook friends). The second part contained two sets of self-reporting items related to the participants’ credibility assessment tasks. Participants were asked to report on the status they had just viewed. In particular, they were asked about three constructs, namely perception of credibility, perception of the text, and level of knowledge in the subject matter.

**Perception of credibility:** a single item was used, “The status is believable” (item 1), ranked on a 6-point Likert scale.

**Perceptions of the text:** we used six items (items 2-7), ranked on a 6-point Likert scale. The main purpose of this category was to validate the assumption of similar texts, that is, to make sure that the two statuses did not differ in their texts (as the participants perceived them). An example item: “The status is informative”. These items were averaged to calculate an index of text perception.

**Level of knowledge in the subject matter:** we used 3 binary (Yes/No) items (items 8-10). The purpose of this set of items was to compare participants’ previous knowledge of the topics presented in the statuses and to validate our assumption of similar levels of knowledge. An example item: “Have you previously heard about the [Assyrian vines/moles’ defecation being able to relieve pain]?”

The full 10 items are presented in Table 2.

The third and last part of the questionnaire, done only after the two tasks of credibility assessment were completed, explicitly mentioned the peripheral information that is usually presented alongside content posted on Facebook and asking the participants about the importance they had attached to each of these components while conducting the credibility assessment task. These components included number of Likes, number of Comments, number of Shares, time stamp, and profile image of the user who posted the status, username of the person who posted the status, the existence of a hyperlink, and the existence of an image. The importance of each component was ranked on a 6-point Likert scale, from 1 (not important at all) to 6 (very important).
Table 2: Part II of the research questionnaire, credibility assessment tasks

<table>
<thead>
<tr>
<th>#</th>
<th>Item/Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>The status is believable</td>
</tr>
<tr>
<td>2</td>
<td>The status is informative</td>
</tr>
<tr>
<td>3</td>
<td>The status is up to date</td>
</tr>
<tr>
<td>4</td>
<td>The information in this status is interesting to me</td>
</tr>
<tr>
<td>5</td>
<td>The status was easy for me to understand</td>
</tr>
<tr>
<td>6</td>
<td>I think the person who wrote the status is an expert in the field</td>
</tr>
<tr>
<td>7</td>
<td>I would like to keep reading and deepening my knowledge of this subject</td>
</tr>
<tr>
<td>8</td>
<td>Have you previously heard about the [Assyrian vines/moles’ defecation being able to relieve pain]? [Yes/No]</td>
</tr>
<tr>
<td>9</td>
<td>Answer this item only if you answered &quot;Yes&quot; in the previous item; otherwise end here. Have you previously heard about the [Assyrian wine/ability to use moles’ defecation in painkillers]? [Yes/No]</td>
</tr>
<tr>
<td>10</td>
<td>Answer this item only if you answered &quot;Yes&quot; in the previous item; otherwise end here. Is the information in the status compatible with your knowledge about it? [Yes/No]</td>
</tr>
</tbody>
</table>

**Procedure**

As mentioned above, participants were recruited while studying in a university library. The researchers (the three authors) were positioned in a closed, quiet room within the library. Participants were recruited individually; each was asked to enter the room separately. After being told what was going to happen, and after they signed an informed consent form to participate in the study, they were seated in front of a computer connected to the eye tracking camera.

Participants filled in the first part of the pen and paper questionnaire (demographic and Facebook use characteristics). Then, they went through a short session to calibrate the eye tracker, after which the first status was presented to them on the computer screen. They were instructed to watch the status as long as they wanted, until they could determine the credibility it assigned to it; their eye gaze was recorded throughout this process using the eye tracking camera. When they finished, the recording was stopped, and the participants were asked to fill in the section in the second part of the questionnaire on the watched status. This process of calibration, status watching, and filling in the questionnaire was repeated for the second status. About half of the participants (arbitrarily chosen) watched the positive status (Assyrian wine) first, and the rest watched the negative status (mole defecation) first. Finally, participants were asked to fill in the third part of the questionnaire. The whole process took about 15 minutes per participant.

It is important to note that only after watching the two statuses and filling in the corresponding parts of the questionnaire — that is, while filling in the third part of the questionnaire, where the peripheral components were explicitly mentioned — were the participants exposed to any mention of the various status components. When initially presenting the research to the participants, we told them the study was about credibility assessment in social networking sites; we did not mention differences between the statuses, nor did we mention possible cues in the statuses. Obviously, we wished to avoid any bias in the participants’ a priori attitudes to either of the statuses. Therefore, the language of the instructions was extremely important; for example, participants were explicitly asked to "watch the status", not to "read" it.

**Data Processing**

To explore the mechanism of viewing the different status components, an *Area of Interest* (AoI) was defined for each one. That is, each component had a corresponding physical area, defined on the two dimensional plane of pixels. Note that these AoIs were non-overlapping; they were used only for data analysis and were not presented during data collection; also note that these AoIs did not fully
cover the whole image, and there were blank spaces between them. See Figure 2 for a demonstration of the AoIs for one of the statuses. We used these AoIs to define the measures of viewing the statuses.

**Figure 2: Areas of interest (AoIs) corresponding to the status components (the status presented here is an English translation of the original status shown to the participants)**

**RESEARCH VARIABLES**

The main variables in the analysis of the mechanism of credibility assessment referred to the location of the various status components, that is, for each Area of Interest (AoI) measures were defined for gaze, fixation and order. For a review of eye tracking metrics, see Poole and Ball (2005).

When analyzing data from an eye tracker, the most basic metrics are related to gazing; a gaze point is the point at which the participant is looking at a given moment. As mentioned above, the eye tracker that was used has a frequency level of 60 Hz, where data is captured 60 times per second. Using this data, it was possible to estimate the overall gaze time for a given image.

Fixation refers to a series of gaze points that are very close in space and time, and the gaze lasts for at least a given time, above a predefined threshold. In this case, based on the specification of the employed technology, fixation was defined as a threshold of 150 milliseconds. Using this metric, *percentage of fixations in an AoI* was calculated. That is, for each AoI, calculating the percentage of fixations on this AoI out of the overall fixations on the entire image.

Using fixation data, it was also possible to measure *time to first fixation on an AoI*, that is, the time that passed from the beginning of the recording to the first fixation on this AoI. The lower the value of this variable, the earlier the participant had viewed that AoI and the higher its priority. For example, for a given participant, a given area yielded a value of 1 for priority if the participant’s earliest fixa-
tion was in this area. (Note that if users never fixate on a given AoI, it will not have a priority value.) This enabled us to calculate the *AoI priority* for each area as the average of its priority across all participants. Finally, fixation data enabled to model *transitions between AoIs*; this is discussed in the Findings section.

Based on the questionnaire, we calculated the *importance of peripheral cues in assessing credibility*, referring to the following cues (each separately): number of lines, number of Comments, number of Shares, status recency, user profile image, username of the person who posted the content, availability of hyperlink, and availability of image.

**Findings**

**Design Validation**

Bearing in mind that the design of the two statuses was such that they were supposed to affect the participants positively or negatively based on the peripheral cues only was necessary to validate that the participants did not perceive the status texts differently. To do so, participants' perceptions of the texts using six items (items 2-7 in the questionnaire) was measured. Cronbach's alpha for these items had a value of 0.7, which is acceptable for a preliminary exploratory study like this one (Peterson, 1994). Therefore, these items were averaged to create a *text perception* index. Item statistics and index statistics are summarized in Table 3.

<table>
<thead>
<tr>
<th>Item (sorted by p value, increasing)</th>
<th>Mean (SD), Positive Status</th>
<th>Mean (SD), Negative Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>2. The status is informative.</td>
<td>4.4 (1.3)</td>
<td>4.3 (1.4)</td>
</tr>
<tr>
<td>3. The status is up to date. †</td>
<td>4.1 (1.3)</td>
<td>4.2 (1.2)</td>
</tr>
<tr>
<td>4. The information in this status is interesting to me.</td>
<td>3.5 (1.5)</td>
<td>3.7 (1.5)</td>
</tr>
<tr>
<td>5. The status was easy for me to understand.</td>
<td>5.0 (0.9)</td>
<td>5.4 (0.6)</td>
</tr>
<tr>
<td>6. I think the person who wrote the status is an expert in the field.</td>
<td>3.0 (1.1)</td>
<td>2.9 (1.6)</td>
</tr>
<tr>
<td>7. I would like to keep reading and deepening my knowledge of this subject.</td>
<td>2.7 (1.5)</td>
<td>3.4 (1.7)</td>
</tr>
<tr>
<td>Average</td>
<td>3.8 (0.8)</td>
<td>4.0 (1.0)</td>
</tr>
</tbody>
</table>

† N=18 for this item;

Despite the small sample size, the standard t test was suitable, as the outcome of normality tests was satisfying. Specifically, we referred to Kim's (2013) guidelines for evaluating normality by calculating

\[ Z_{\text{Skewness}} = \frac{\text{Skewness}}{SE_{\text{Skewness}}} \quad \text{and} \quad Z_{\text{Kurtosis}} = \frac{\text{Kurtosis}}{SE_{\text{Kurtosis}}} \]

if either of these (absolute) values is larger than 1.96 (for samples smaller than 50), the distribution should be considered non normal. For the positive status, \( Z_{\text{Skewness}} \) and \( Z_{\text{Kurtosis}} \) yielded values of 1.24 and 1.01, respectively; for the negative status, \( Z_{\text{Skewness}} \) and \( Z_{\text{Kurtosis}} \) yielded values of 1.05 and 1.18, respectively. Therefore, we assumed normality and continued with the standard t test. When comparing these two variables, no significant difference between them was found, with \( t(18)=0.73 \), at \( p=0.47 \). Thus concluding that the statuses' text similarity was validated.

**Status Design and Perceived Credibility**

Next, the item measuring credibility (item 1) was approached. Its skewness values for the positive and negative statuses were -0.66 and -0.52, respectively, with SE=0.52, thus giving Z values of 1.26 and
0.98, respectively. Kurtosis values for the positive and negative statuses were -0.78 and -0.63, respectively, with SE=1.01, giving Z-values of 0.77 and 0.62, respectively. The four Z-values were below Kim's (2013) threshold; therefore, assuming normality.

*Perceived credibility* had a mean of 4.16 (SD=1.30) for the positive status, and 3.16 (SD=1.26) for the negative status. This difference was significant, with t(18)=2.21, at p<0.05. Therefore, it was concluded that there was a connection between the status design and its perceived credibility.

## Gazes and Fixations

There was no significant difference in the overall time it took to watch each of the statuses. On average, the positive status took 32.8 seconds (SD=13.1), and the negative status took 33.9 seconds (SD=12.0). Skewness values for the positive and negative statuses were 1.85 and 0.97, respectively, with SE=0.52, thus giving Z values of 3.52 and 1.85, respectively. Kurtosis values for the positive and negative statuses were 4.40 and 0.45, respectively, with SE=1.01, giving Z values of 4.34 and 0.44, respectively. Only one of these four Z values was far below Kim's (2013) threshold; the others were either larger or very close to it. Therefore, it was not possible to assume normality, and the Wilcoxon Signed Rank test was used to compare the gazing data. The test resulted in Z=-0.16, at p=0.87, leading to conclude that the viewing times for the two statuses were not significantly different.

Throughout all participants' sessions, the eye tracking system recorded 1534 fixations on the positive status (Assyrian wine) and 1573 fixations on the negative status (mole defecation). Area 8 (indications of Comments) had only 6 fixations; area 9 (indication of Shares) was absent in the negative status and had no fixations on the positive status, so we omitted this AoIs' data. Fixations that occurred outside the AoIs were also omitted. This reduced the data to 1192 fixations on the positive status and 1198 fixations on the negative status.

The average number of fixations per participant was 62.7 (SD=16.9) for the positive status and 63.1 (SD=22.1) for the negative status. Skewness values for the positive and negative statuses were -0.14 and 0.60, respectively, with SE=0.52, giving Z values of 0.27 and 1.15, respectively. Kurtosis values for the positive and negative were 1.31 and -0.73, respectively, with SE=1.01, giving Z values of 1.29 and 0.72, respectively. As these four Z values were much lower than Kim's (2013) threshold, normality was assumed, and the two means were compared using the t test, resulting in t(18)=0.08, at p=0.94. Hence, the average number of fixations on the two statuses was not significantly different.

Not surprisingly, the area with the most fixations in both statuses was the text area (area 4), with 1007 fixations (85%) on the positive status and 931 (78%) on the negative status. Notably, in both cases, the second most popular area was the image area (area 6), with 127 fixations (11%) on the positive status and 192 fixations (16%) on the negative status. The third most popular area was the hyperlink area (area 5), with 25 fixations (2%) on the positive status and 24 fixations (2%) on the negative status. The remaining areas, i.e., profile image (area 1), username (area 2), posting timestamp (area 3), and indications of Likes (area 7), received less than 2% of the fixations in each of the statuses. Results are summarized in Table 4. When comparing the distributions of fixations for the two statuses, a significant difference was found, $\chi^2(6) = 26.2$, at p<0.001.

Bearing in mind that for AoI priorities, the higher the number, the higher the priority of the relevant AoI, some interesting differences were found in the AoI priorities of the two statuses when comparing mean values across all participants. On average, in both cases, the highest priority (first fixation) was the image (area 6), and the second most prioritized area was the text (area 4). After these, the two statuses diverged.

For the positive status, the other AoIs, sorted by their AoI priority, were: profile image, username, posting timestamp, indication of Likes (with the same value), and hyperlink. For the negative status, the other AoIs, sorted by their AoI priority, were hyperlink, profile image, username, indication of Likes, and posting timestamp.
Table 4. Number (and %) of fixations on each Area of Interest (AoI)

<table>
<thead>
<tr>
<th>AoIs</th>
<th>Number of Fixations (%)</th>
<th>Positive Status (Assyrian Wine)</th>
<th>Negative Status (Mole Defecation)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 – Profile Image</td>
<td>6 (0.5%)</td>
<td>19 (1.6%)</td>
<td></td>
</tr>
<tr>
<td>2 – Username</td>
<td>10 (0.8%)</td>
<td>13 (1.1%)</td>
<td></td>
</tr>
<tr>
<td>3 – Posting Timestamp</td>
<td>9 (0.8%)</td>
<td>15 (1.3%)</td>
<td></td>
</tr>
<tr>
<td>4 – Text</td>
<td>1007 (84.5%)</td>
<td>931 (77.7%)</td>
<td></td>
</tr>
<tr>
<td>5 – Hyperlink</td>
<td>25 (2.1%)</td>
<td>24 (2.0%)</td>
<td></td>
</tr>
<tr>
<td>6 – Leading Image</td>
<td>127 (10.7%)</td>
<td>192 (16.0%)</td>
<td></td>
</tr>
<tr>
<td>7 – Indication of Likes</td>
<td>8 (0.7%)</td>
<td>4 (0.3%)</td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>1192 (100%)</strong></td>
<td><strong>1198 (100%)</strong></td>
<td></td>
</tr>
</tbody>
</table>

**Reported Importance of Peripheral Cues**

Using the data collected in the questionnaires, the level of importance that the participants attached to each of the peripheral cues (N=19) was measured. Of highest importance was the hyperlink (M=4.84, SD=1.61), followed by the image (M=3.63, SD=1.95), recency (M=3.58, SD=1.87), username (M=3.47, SD=1.58), number of Comments (M=3.05, SD=1.72), number of Shares (M=2.89, SD=1.79), and user profile picture (M=2.84, SD=1.95). Of least importance, based on the participants' self-reports, was the number of Likes (M=2.63, SD=1.54).

**Transitions Between Areas**

To better understand the mechanisms involved in examining a status before making a decision on its credibility, the transitions of fixations between pairs of AoIs were analyzed. From these transitions, a directed transition graph was built, modeling the probability of moving from one AoI to another. The graph is shown in Figure 3; each node indicates an AoI, and the (directed) edges between nodes denote transitions between the corresponding AoIs.

**Aggregated exploration**

First, analyzing transitions at the transition level—that is, considering any transition made by any of the participants during the session—was conducted. Overall, there were 169 transitions between AoIs for the positive status (Assyrian wine), and 205 transitions between AoIs for the negative status (mole defecation). For purposes of simplicity, only transitions that occurred at least four times are discussed. In the transition graphs shown in Figure 3, the numbers on each edge represent the number of corresponding transitions, and the percentage of transitions is calculated from the source node. That is, the percentage values on each node's outgoing edges should sum up to 100 percent. However, for the findings to be more meaningful, we only present the most common transitions.

The most common transitions in the positive status (Assyrian wine), in absolute values, were from the image to the text (40 transitions) and from the text to the image (35 transitions). The next two most common transitions were from the text to the hyperlink (18 transitions) and from the hyperlink to the text (15 transitions). The other transitions occurred 10 times or less. The most popular ingoing AoI was the text area (74 ingoing transitions), followed by the image area (44 transitions), hyperlink (21 transitions), user image and timestamp (9 transitions each), number of Likes (7 transitions), and user image (only 5 ingoing transitions).

The most common transitions in the negative status (mole defecation), in absolute values, were from the image to the text (45 transitions) and from the text to the image (45 transitions); the next two most common transitions were from the user image to the text (14 transitions) and from the hyperlink to the image (11 transitions). The remaining transitions occurred 10 times or less. The most popular ingoing AoI was the text area (with 80 ingoing transitions), followed by the image area (61
transitions), then user image (18 ingoing transitions), hyperlink (15 transitions), timestamp (14 transitions), username (13 transitions), and number of Likes (only 4 ingoing transitions).

These findings are summarized in Figure 3. When looking at the two transition graphs, it is possible to observe some additional differences between the positive (top) and the negative (bottom) cases, most notably the absence in the positive status of transitions between the user image and the text AoIs (in both directions), as well as a lack of transitions from the hyperlink to the images and from the timestamp to the user image. The negative status lacks transitions from the text to the number of Likes, but these appear for the positive status.

Figure 3: Transition graphs for positive (left) and negative (right) statuses; representing all transitions observed at least four times. Percentage is for all outgoing transitions (including those not represented)

Participant level analysis
When analyzing transitions at the participant level, each transition observed in a participant's data was counted only once for this participant. Taking into consideration transitions observed for at least four participants enabled to explore common transitions.

Figure 4 shows the findings. In this case, numbers on each edge (in percentage) represent the frequency of the corresponding transition (N=19 equals 100%). Overall, eight unique transitions were observed for the positive status and 12 for the negative status. In both cases, the most dominant AoI was the text area; it served as a hub for most transitions (for all transitions, in the positive case). In both statuses, transitions between the text and the image areas (in both directions) were the most common and were observed for at least 16 participants (84%). In the positive status, the second most common pair of AoIs in which we observed transitions was the text and hyperlink areas; each transition was observed for at least eight participants (42%). In the negative status, the most common pair was the text and user image areas, with each transition was observed for at least eight participants (42%).

Interestingly, all AoIs between which transitions were observed for the positive status were also found for the negative status. However, two pairs of AoIs were observed for the negative status but
not for the positive one: text and user image, and image and hyperlink. These findings are summarized in Figure 4.

![Transition graphs for the positive (left) and the negative (right) statuses; multiple transitions between AoIs were counted only once for each participant, and appear here only if they were observed for at least four participants. Percentage represents all participants (N=19)](image)

**Participant level distinguishing analysis**

To deepen the understanding of the differences between the two scenarios—still at the participant level—only those transitions observed at least once for a given participant in one status and not observed for the same participant in the other status were explored. As in the previous subsection, only common cases were considered, i.e., if they occurred for at least four participants.

When comparing positive and negative status, a different behavior in four transitions was found. Six participants (32%) transitioned from the text to the hyperlink AoIs or vice versa in the positive status, but not in the negative status. The same was true for the text and username AoIs, albeit to a lesser extent (text to username, five participants, or 26%; username to text, four participants, or 21%).

In contrast, eight transitions appeared in the negative status but not in the positive status. At the former case, four participants (21%) transitioned from the text and the username AoIs or vice versa in the negative status but not in the positive status. In addition, four and five participants (21% and 26% respectively) transitioned from the text to the timestamp or vice versa in the negative but not in the positive status. Also, and not surprisingly — since they were not prominent at all in the positive case — transitions between the user image and the text AoIs, as well as between the image and the hyperlink AoIs (in either direction), were observed for a few participants in the negative status and not in the positive case.
Figure 5: Transition graphs for positive (left) and negative (right) statuses; represented transitions that were observed (by the same participants) in one status and not in the other, if this occurred for at least four participants. Percentage represents all participants (N=19)

**SUMMARY OF FINDINGS**

Overall, a connection was found between the status design and its perceived credibility, with the positively framed status (Assyrian wine), compared to the negatively framed (mole defecation), perceived as more credible.

Using an eye tracking camera, it was possible to examine the ways in which the participants had examined both statuses. It was found that there were no differences in the overall time taken to examine both statuses, and that the three most viewed AoIs in both cases (counting number of fixations) were the text, the leading image, and the hyperlink. This is in line with the participants' self-report, based on which the most important message cues for credibility are hyperlink and image.

Analyzing the transitions between AoIs, it was found that in both statuses, the most common transitions were from the image to the text and from the text to the image. In the positive status, the next two common transitions were from the text to the hyperlink and from the hyperlink to the text, while in the negative status they were from the user image to the text and from the hyperlink to the image. Additionally, the positive status lacked transitions between the user image and the text AoIs (in both directions), as well as transitions from the hyperlink to the images and from the timestamp to the user image; The negative status lacked transitions from the text to the number of Likes, while these showed up for the positive status.

Analyzing transitions at the participant level, it was found that the most dominant AoI in both statuses was the text area; it served as a hub for most transitions (for all transitions, in the positive case). In both statuses, transitions between the text and the image areas (in both directions) were the most common. In the positive status, the second most common transition was between the text and hyperlink areas, and in the negative status it was between the text and user image. Interestingly, all AoIs between which we observed transitions for the positive status were also found for the negative status; however, two pairs of AoIs were observed for the negative status but not for the positive one: text and user image, and image and hyperlink.
Still at the participant level, looking specifically at transitions that were unique to one status, we found transitions between the text and the hyperlink AoIs, as well as between the text and the username AoIs, which appeared in the positive status but not in the negative status. In contrast, transitions between the text and the timestamp, as well as between the text and the username, appeared in the negative status but not in the positive status.

**DISCUSSION**

This study explored the mechanisms of credibility assessment of scientific information posted on Facebook. Given the richness and importance of peripheral cues on social networking sites (e.g., user profile image, number of Likes, etc.), we were interested in their role in the assessment process. More specifically, a within subject approach was taken to explore differences between positive and negative framing using these peripheral cues. The analysis of the fine-grained data collected using an eye tracking camera shed a new and interesting light on the effects of this framing on credibility assessment processes.

The positively framed status was perceived, as expected from the Elaboration Likelihood Model (ELM), as more credible than the negatively framed status, with no differences in the readability or understandability of the text itself. This clearly demonstrates the effects of the visual framing, a finding in line with previous studies of the role of framing in credibility or trust assessment (Harris, Sillence, & Briggs, 2009; O’Keefe & Jensen, 2008; Pengnate & Antonenko, 2013). In addition, positive/negative affect was associated with the positive/negative framings, a finding in line with previous studies of affect in information processing (Fang, 2014; Kim, Kiousis, & Molleda, 2015). As the findings indicate, the differences in participants’ mechanisms of assessing credibility between the two scenarios were not evident in the aggregated viewing behavior. They were evident, however, in the specific ways participants examined the various status components. Note that this more nuanced understanding was enabled by the unique methodology taken in this study.

Overall, the text and the image areas were the most prominent in both assessment processes. The text prominence was obvious and not surprising, as users were explicitly guided to assess the credibility of the presented statuses, hence were drawn to the text area which was prominent in its size. The image prominence was not surprising either; previous studies note the important role visual aids play (Fogg et al., 2003; Metzger, Flanagan, & Medders, 2010). Third in order of prominence was the hyperlink area; this may indicate the participants’ examination of the information source rather than the information sharer. This distinction is particularly important in social networking sites (SNS), as information may be reposted many times, creating layers of information sharers and obscuring the original source. As such, the findings are in line with previous studies indicating the importance of hyperlinks in credibility assessment (Borah, 2014; Johnson & Wiedenbeck, 2009). Moreover, hyperlinks and images were the two most important factors in that process, as reported by the participants. That is, both the eye tracking and self-report data highlight the importance of these two peripheral cues.

As mentioned above, the methodology used here permitted a nuanced understanding of the credibility assessment process. By examining the transitions between various areas on the screen, it was possible to observe how the mechanism differed when the overall context was positive or negative. The negative status had 20% more transitions than the positive status, indicating more profound (but not longer) information evaluation. Transitions were found from the text to the area indicating the number of Likes only for the positive status; this resonates with previous findings, according to which popularity is more influential when the framing is positive (Borah & Xiao, 2018).

When examining these transitions at the participant level, some clear within subject differences were found between the positive and the negative statuses. In the negative case, we observed transitions that were not present in the positive case, specifically transitions to or from the user image, as well as transitions between the image and the hyperlink. That is, the importance of the source — be it the
information sharer or its original source—grows when the framing is negative; this finding supports previous explorations of the interaction between source and framing (Creyer, 1997; Hussein, Manna, & Cohen, 2014; Kim & Kim, 2014). Moreover, as both out of the network and within the network source related cues were taken into consideration (the former by adding the hyperlink, the latter by adding the user image and the username), it is argued that adult users of SNS do acknowledge that on these platforms the content is shared by various sources, and hence may represent a proprietor’s content or a message from network peers (Neubaum & Krämer, 2017). However, as recent studies have found no significant effects of profile image related features (e.g., gender, facial expression) on credibility assessment (Wang, 2016; Xu, 2014), more research is required to determine the specific role of the profile image.

**CONCLUSIONS**

This study contributes to the growing literature on credibility assessment in SNS by considering the assessment process rather than just focusing on the resulting perceived credibility. When people admit posting fake information on social networking sites (Buzzetto-More, Johnson, & Ellobaid, 2015), and when these platforms are used not only for entertainment, but may also serve for professional development (Wandera, James-Waldon, Bromley, & Henry, 2016), the findings of this study have some farfetching implications. It was recently shown that digital literacy and Internet experience are negatively associated with credibility assessment of online materials, hence it makes Internet users more critical about the content they encounter online (Shen et al., 2019). Indeed, it is the goal of many media literacy programs worldwide to evaluate incoming information (Manzoor, 2018), and evidence exists of the explicit benefit of such programs in the specific context of online information verification (Seo, Erba, Altschwager, & Geana, 2019). As part of digital literacy education, the major focus should be given to the role of peripheral cues on credibility assessment in social networking sites. Educators should emphasize the mechanisms by which these cues interact with message framing, so Internet users would be encouraged to reflect upon their own credibility assessment skills, and eventually to improve them.

The findings on the mechanisms of that process, enabled by the use of eye tracking data, point to the different roles of specific peripheral cues, when the message is overall peripherally positive or negative. Thus, this study also contributes to the theoretical literature on framing effects in science communication, as we highlight the peripheral cues that make a strong frame. This understanding leads the way to a new set of research questions emphasizing the process of credibility evaluation.

This study is not without limitations. As it represents a relatively small population size, the results of the statistical comparisons should be considered cautiously. In addition, the items the participants were asked to evaluate were fabricated. Note that this is a common practice in assessing perceptions of SNS content, specifically credibility (e.g., Turcotte, York, Irving, Scholl, & Pingree, 2015). Finally, the research population, recruited at a single university campus in Israel, was not necessarily representative of relevant generalized populations, e.g., the nation population, or undergraduates in the country. Still, we believe that the current study makes an important contribution to the understanding of credibility assessment processes in the digital age.

Addressing these issues, we recommend that future studies should consider data collection in a more authentic context (using, for example, a think aloud protocol while participants browse their own Facebook page), referring to more SNSs, and studying additional populations (which will vary by, e.g., age, education level, and cultural aspects).

**REFERENCES**

Eye-Tracking Study of Credibility of Scientific Information


Eye-Tracking Study of Credibility of Scientific Information


Eye-Tracking Study of Credibility of Scientific Information


**Biographies**

**Aviad Rotboim** is a Ph.D. student in Tel Aviv University's School of Education, Israel. His research is focused on credibility assessment of scientific information in social networking sites. He holds an M.A. in Biostatistics and a B.A. in Statistics and Economics. He is a lecturer in several academic institutions, teaching statistics, probability and data analysis.
Arnon Hershkovitz, Ph.D., is a Senior Lecturer in the School of Education, Tel Aviv University, Israel. His research lies in the intersection of learning, teaching, and technology. Mostly, he is interested in the skills requested for, and shaped by, today’s settings of learning and teaching, which are part of a cyberlearning ecosystem; these are studied using Learning Analytics and other methodologies.

Eddie Laventman, M.A., is a former graduate student in the School of Education, Tel Aviv University, Israel. He holds a B.A. in Human Resources. He is an international expert and keynote speaker on social media, storytelling, and marketing in the digital age. He serves as a consultant to the biggest brands in Israel and worldwide. He is a senior lecturer and course manager in a variety of MBA programs and professional courses at Israel’s leading academic and professional institutions.