City, University of London Institutional Repository

Believing in Credibility Measures: Reviewing Credibility Measures in Media Research From 1951 to 2018

ANINA HANIMANN[1]
University of Lucerne, Switzerland

ANDRI HEIMANN
Centre for Democracy Studies Aarau, Switzerland

LEA HELLMUELLER
City, University of London, UK

DAMIAN TRILLING
University of Amsterdam, The Netherlands

Although credibility has been a key concept in communication research for decades, there still is no consensus on its conceptualization and measurement. Indeed, scholars have criticized the lack of theory-driven approaches, conceptual inconsistencies between sub-constructs of credibility, and the problems of applying them to the contemporary media environment. This literature review of quantitative studies of credibility published between 1951 and 2018 explores state-of-the-art definitions and measures of credibility (N = 259). While most studies make a conceptual distinction between source, media, and message credibility, measurement scales do not follow this traditional trinity. Instead, we propose moving toward a dual-credibility model.

Keywords: media credibility, source credibility, message credibility, measurement

The degree to which journalists, politicians, and other actors succeed in reaching their audiences through media messages highly depends on citizens' perceptions of credibility (Golan, 2010). The credibility of the provided information can influence information selection (Winter & Krämer, 2014),

attitude change (Lupia, 2000; Petty & Cacioppo, 1981), or behavioral responses (Iyengar & Valentino, 2000). Credibility has therefore received more scholarly attention than many other variables related to communication processes (Kiousis, 2001). Furthermore, recent discussions about the advent of a post-factual era and the spread of both misinformation and disinformation have cast renewed attention on credibility.

However, scholars have identified several conceptual and methodological shortcomings of existing research on credibility. For example, studies often fail to differentiate credibility from related concepts, such as trust. Some see trust as an antecedent to credibility (e.g., Hong, 2006), others as the result of a longer process in which credibility is but a single factor (e.g., Rowley, Johnson, & Sbaffi, 2015), and still others as synonymous with credibility (e.g., Tsfati & Cappella, 2003). Building on previous research, we make a clear distinction between trust and credibility: Trust is an attitude that rests on prior experience, while credibility is a situational evaluation that involves some form of information processing (Go, You, Jung, & Shim, 2016) and is thus limited to the communicative dimension (Bentele & Seidenglanz, 2008).

Scholars have also identified the lack of a clear theoretical definition of the concept of credibility as a root cause for problems related to its measurement (McCroskey & Young, 1979). Scholars started adopting an empirical approach regarding this problem of insufficient conceptual definitions as early as 1951. These early studies sought to identify the relevant dimensions of credibility through factor analyses. However, many of them failed to provide comprehensive construct definitions and remained heavily data driven (Kohring & Matthes, 2007).

After nearly 70 years of research on credibility it is time to take stock of what we have, to identify ways of moving forward. Outlining how the heritage of research on credibility can inform current studies of the changing media landscape and of mis- and disinformation seems paramount. Therefore, this study provides an overview of the state of the art in conceptualizing and measuring credibility for the purpose of guiding future quantitative studies on credibility.

The Construct of Credibility in Research on Mediated Communication

Within the field of mediated communication, credibility is often defined as "a perceiver's assessment of believability or of whether a given speaker is likely to provide messages that will be reliable guides to belief and behavior" (Simons, 2002, p. 20). Such a receiver-oriented conceptualization of credibility has dominated research within psychology (Hovland, Jannis, & Kelly, 1953), communication science (Metzger, Flanagin, Eyal, Lemus, & McCann, 2003), and marketing research (Eisend, 2003). Accordingly, credibility is conceptualized as a quality receivers ascribe to people, institutions, or their communicative products (Bentele & Seidenglanz, 2008).

Source Credibility

Early studies of credibility focused on source credibility defined as a two-dimensional construct consisting of expertise (the source can make valid assertions) and trustworthiness (confidence in the

honesty of a source; Hovland et al., 1953). While these two dimensions are still widely used, scholars who adopt approaches based on factor analyses make a case for including additional dimensions, such as dynamism (the energy available to a source; Berlo, Lemert, & Mertz, 1969) or attractiveness (physical attractiveness; Ohanian, 1990). Another potential third dimension, goodwill (the source's good intent toward society), dates back to Aristotle's early conceptualization of source credibility (ethos; McCroskey & Teven, 1999). Nevertheless, some scholars consider the latter a part of the trustworthiness dimension (Peters, 1992). The focus on different dimensions of source credibility can result from both the specific construct under investigation and the receivers' needs (Berdahl, Bourassa, Bell, & Fried, 2016): Research on source credibility moved from investigating the speakers' credibility (Hovland et al., 1953) to examining the credibility of media organizations (Fico, Richardson, & Edwards, 2004). Therefore, it seems plausible that the credibility of an organization relates to individuals' credibility at a conceptual level (i.e., both are *sources* of information) but differs from it regarding its underlying dimensions (e.g., attractiveness as a dimension is important for individuals but less so for organizations).

Some scholars perceive the variety of identified dimensions as a direct result of studies' lack of theory and their heavy focus on measurement (Metzger et al., 2003). They have, for example, criticized that these dimensions capture individuals' more general perceptions of the source (McCroskey & Young, 1981) or that the dimensions are correlates, rather than parts, of credibility (Kohring & Matthes, 2007).

Media Credibility

Scholarly interest in the concept of media credibility was initiated by the Roper Organization in the United States. Its studies asked people about the medium they would most likely believe if they received conflicting reports (Roper Organization, 1985). The Roper studies' methodological approach has been criticized on multiple fronts: The question wording may be biased, people may use different reference points to evaluate different media, and the reliance on a single indicator may be insufficient (for an overview, see Kohring & Matthes, 2007). The locus of evaluation also varies, depending on what media people evaluate: Individuals (e.g., news anchors) in the case of TV and institutions in the case of newspapers (Newhagen & Nass, 1989). Consequently, similar to their colleagues working on source credibility, scholars of media credibility used factor analytical approaches to derive media credibility's underlying dimensions. To do so, they relied heavily on items used to measure source credibility, which raised concerns about a conceptual overlap between the two constructs (Metzger et al., 2003).

To respond to the need for a more specific measure of media credibility, Gaziano and McGrath (1986) asked people to evaluate the daily newspaper and TV news they were most familiar with on a series of items. They derived a single-factor solution that consisted of items such as perceptions of fairness or accuracy. Meyer (1988) further developed this work, and West (1994) then tried to derive a standardized scale meant to improve cross-study comparisons. He concluded that Meyer's (1988) five-item scale based on fairness, bias, comprehensiveness, accuracy, and trustworthiness constituted a valid and reliable measure of media credibility. Yet, like most other authors, West (1994) did not provide a theoretical definition of media credibility. Therefore, some crucial questions about how media credibility is conceptualized remain unanswered (Kohring & Matthes, 2007).

Message Credibility

Unlike source and media credibility, message credibility has received little attention as a dependent variable and is mainly studied as an independent variable (Metzger et al., 2003). Thorough conceptualizations of message credibility are therefore scarce. Yet, scholars argued early on that message credibility (i.e., the credibility of a message's content) is theoretically distinct from source credibility, and when information about the source of a message is absent, people assess its credibility based on its content (Rosenthal, 1971). Recent contributions have shown that source and message credibility are interrelated (Schweiger, 2000).

Message credibility has been defined as a perception of information accuracy, reliability, validity, and objectivity (Metzger et al., 2003; Sundar, 1999). Even though attempts to operationalize message credibility are rare, some exist. Smith (1978) made an early contribution by developing a measurement inventory designed to capture message credibility. However, because his operationalization of message credibility is based on items measuring source credibility, it confounds source and message credibility (Metzger et al., 2003). Sundar (1999) proposed a three-item scale meant to measure the credibility of news stories. The latter was never empirically validated. Some scholars (e.g., Meyer, Marchionni, & Thorson, 2010; Pjesivac, Geidner, & Cameron, 2018) rely on Meyer (1988) and Gaziano and McGrath (1986) to measure message credibility. Nevertheless, these conceptualizations are not completely distinct from source credibility because of the latter's focus on media credibility.

More recently, communication scholars have begun to develop more message-based conceptualizations. For instance, Appelman and Sundar (2016) developed and validated a three-item scale (i.e., accurate, authentic, and believable) to assess the credibility of news.[2]

The Trinity of Source, Media, and Message Credibility—An Outdated Approach?

Since the turn of the new millennium, scholars have repeatedly questioned how traditional credibility measures fit the then-new Web-based communication context (Metzger et al., 2003). For instance, Borah (2014) indicates that "the introduction of any new medium should make us rethink the basic theoretical assumptions in our field" (p. 576). This raises the question whether the distinction between source, media, and message credibility can be upheld in the contemporary media environment.

Crucially, in Web-based communication, it is no longer self-evident what the source is. Individuals may perceive the source to be the website operator, the medium, or the writer of an online article, or even consider all these levels when asked about a source's credibility (Metzger et al., 2003; Schweiger, 2000; Winter & Krämer, 2014). While this problem is arguably inherent to all mediated communication, online communication aggravates it as users can easily share information via diverse channels. Moreover, individuals increasingly consume single news items that are shared not only via mass media but also via personal networks, a phenomenon that Trilling (2019) calls the unbundling of news. Who is considered the source of a news item from a newspaper, which is then shared through

---

[2] Note that this is not the scale developed by Sundar (1999).

personal contact? The same holds for media credibility. What are respondents evaluating when asked about the medium credibility of a newspaper article on Facebook? Does it even make sense to investigate the credibility of the Internet as a medium? For example, Johnson and Kaye (2016) propose to isolate the latter by speaking about genre credibility, and by differentiating among distinct genres, such as blogs or social networking sites.

The attribution of news items to discrete sources and media must be understood within the context of new modes of news dissemination and exposure to news. Trilling (2019) suggests that we should no longer attribute news items to one specific source or medium but instead speak of relationships: One news item can have various relationships with different sources, such as the newspaper the item was published in, or the individual who retweeted the item.

How do these developments affect the way credibility is measured? Existing research has advanced two approaches. In the first, scholars are very explicit about what they seek to investigate (e.g., website operator) and use measures that account for the online context (Metzger et al., 2003; Winter & Krämer, 2014). Nevertheless, this first approach risks ignoring the nested structure of online news consumption and exposure. The second approach encompasses studies that measure several different credibility constructs (e.g., source and message) simultaneously (Kim, 2015). While this approach is more likely to account for the complex structure of online news, it may result in unfeasible measurement instruments.

We may conclude that the traditional concepts of source, media, and message credibility still provide useful starting points, but they are not always capable of addressing the questions that arise when we investigate credibility in the online environment.

Research Questions

This outline shows that even though research on credibility has enjoyed considerable attention, it still faces significant conceptual and methodological challenges, especially in an online and global environment wherein the three constructs of credibility (source, message, and media) seem to multiply (e.g., more than one source) and occur simultaneously. Almost two decades ago, Metzger and colleagues (2003) outlined how credibility research can inform studies of the contemporary media environment. However, their article focused mainly on questions of conceptualization and not measurement. Moreover, the review was published in 2003—the early days of research on credibility in the online context. We, therefore, lack an up-to-date review to guide quantitative credibility studies. Our article seeks to fill this gap by providing a comprehensive overview of how scholars define and measure credibility constructs in both the online and offline contexts. Our primary aim is not to critically analyze these conceptualizations and measurements but to compile and categorize previous credibility research. The derived overview will then serve communication scholars as reference work to make informed choices about the relevant concepts and measurement instruments for their own work. Analyzing the communication and information sciences literature published between 1951 and 2018, we ask the following research questions:

RQ1a: *How do scholars define the three different credibility constructs (source, media, and message credibility)?*

RQ1b: *Do these definitions change over time?*

RQ2a: *What scales and items do scholars use to measure source, media, and message credibility?*

RQ2b: *Do the measurement scales vary depending on the communication context (offline and online)?*

Method

A descriptive literature review is the review method best suited to our purposes of outlining the state of the art in defining and measuring credibility and identifying the patterns and trends that have marked research on the topic (Paré, Trudel, Jaana, & Kitsiou, 2015; Sylvester, Tate, & Johnstone, 2013). Descriptive reviews start with a structured literature search, proceed with identified studies' screening for inclusion or exclusion, and conclude with the extraction and quantitative analysis of relevant characteristics of each study. The following sections outline the procedure in more detail.

Literature Search

We conducted a search of relevant articles published between 1951 and 2018 using Ebscohost and Web of Science databases. We searched for relevant keywords in the articles' titles, abstracts, and full texts. We opted for a broad search strategy and proceeded to manually filter the results (Table A1 in the Appendix).[3] The search resulted in 6,851 hits. Eliminating items that did not appear in the communication journals that the Communication Abstracts database (Ebscohost, 2018) classifies as core or priority journals reduced this number to 1,216 articles (for a similar procedure, see Wallander, 2009). We then applied four criteria to the abstracts to identify the relevant articles. First, the articles had to employ a quantitative approach and measure credibility using a scale (single-item scales were included). Second, credibility had to be the dependent variable.[4] Third, the measured construct had to be source, media, or message credibility.[5] Fourth, the focus had to be on a mediated communication context. Applying these criteria resulted in 227 articles. When the full texts were screened, another 46 articles failed to meet at least one of the four criteria and were excluded, resulting in a tally of 181 articles.

Codebook

Table 1 lists all the variables. Construct category is one of the most important variables we coded. We followed Schweiger (2000) and coded author(s) interested in the credibility of different media types (e.g., the Internet) or subsystems of media types (e.g., blogs in general, online news in general) in the media category. Source was made up of author(s) who investigated the credibility of specific media products

---

[3] Appendix available here: https://doi.org/10.6084/m9.figshare.21443736

[4] Such studies should conceptualize and operationalize credibility in the best possible way.

[5] Articles on the credibility of services/technologies were not included.

(e.g., a specific website), media organizations (e.g., Google), actors (e.g., a president in a newspaper interview), or presenters (e.g., the anchor of a TV show). Finally, message included author(s) who studied the credibility of editorial units (e.g., news items, articles; cf. Table A2 in the Appendix).

Table 1. Research Questions and Coded Variables.

Research Question	Variables
RQ1: Definitions	1. Construct name (e.g., article credibility)
	2. Construct category (i.e., source, message, or media)
	3. Construct details (e.g., individual or collective source)
	4. Communication context (i.e., offline vs. online)
	5. Definition
	6. Dimensionality
	7. Independent variables
RQ2: Measures	8. Scale origin (i.e., [adapted] replication, own scale)
	9. Measurement references
	10. Measurement type (e.g., semantic differential)
	11. Measurement language
	12. Country of study
	13. Measurement length (i.e., number of items)
	14. Response format (i.e., answer categories)
	15. Items
	16. Sample size
	17. Sample type (e.g., student sample)
	18. Method (e.g., survey, experiment)
	19. Validity tests (i.e., conducted or not)
	20. In case of adapted or own scale: Pretest (i.e., yes, or no)
	21. Reliability (i.e., Cronbach's α and inter-item correlation)

After several rounds of adjusting the codebook, Krippendorf's α measuring the inter-coder reliability of the two coders surpassed 0.80 for all coded variables (Freelon, 2010).[6]

Results

We rely on frequency and cluster analyses to produce quantitative results (Paré et al., 2015).

Sample Description

The vast majority of the 181 articles used U.S. data (n = 138). Only 20 articles used data from other countries including Germany (n = 8), China (n = 8), South Korea (n = 5), Switzerland, The Netherlands, Bangladesh, Spain, and the United Kingdom (n = 2 each). Four articles included data from

---

[6] We used roughly 10% of the analyzed scales for calculating the inter-coder reliability (34 of 259). Our data set contained almost no observations for the variable pretest, so we removed this variable from the inter-coder reliability analysis.

more than one country (Fuoli & Hart, 2018; Lock & Seele, 2017; Pjesivac & Rui, 2014; Seo & Lim, 2010). The most prominent journals publishing articles on credibility were *Journalism & Mass Communication Quarterly* (n = 26), *Computers in Human Behavior* (n = 23), *Journal of Computer-Mediated Communication* (n = 10), *Public Relations Review* (n = 9), and *Mass Communication and Society* (n = 8). However, the 181 coded articles appeared in a total of 66 different journals. Scholars have become more interested in credibility, especially since 2010 (Figure 1).

Most articles either relied on survey-embedded experiments (n = 88) or on regular surveys (n = 62). Twenty-one articles used laboratory experiments, and seven articles combined different methods (e.g., to develop measurement scales). Only two articles used content analysis.[7,8] There was a huge variation in the articles' sample sizes, ranging from 20 to 6,738 respondents (Median = 262). Most of the articles that employed regular surveys (Median = 442, SD = 955.1) and surveys with experiments (Median = 228.5, SD = 743.7) used samples that were moderate in size. However, large-scale surveys (N ≥ 1,000) remained scarce (n = 19). This may stem from the type of the used sample: Almost half of the articles rely on students (n = 88) or other convenience samples (n = 53). Only 16 articles use random citizen samples.
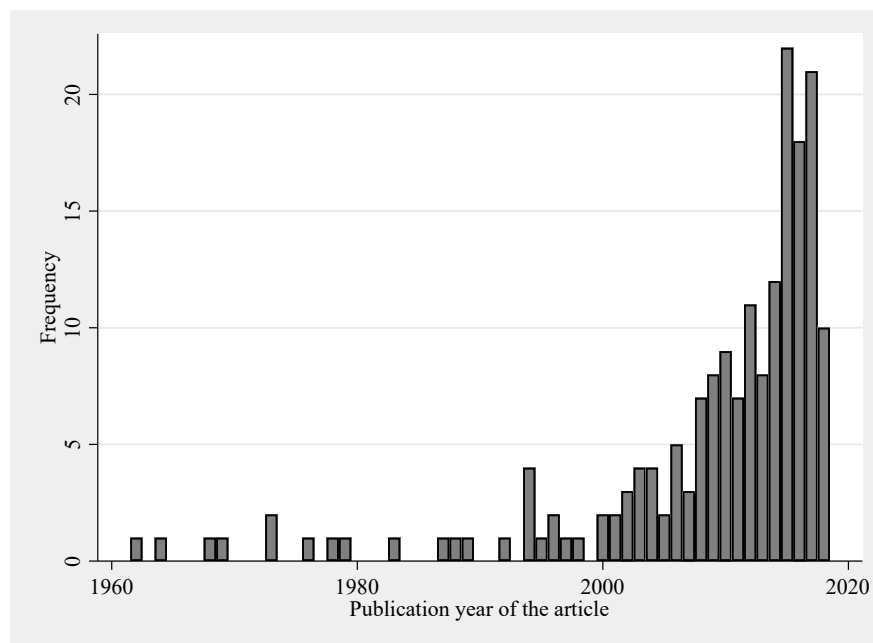
*Figure 1. Publication year of the articles in the sample.*

---

[7] In this case, coders rated the credibility scales.

[8] We could not determine the method used by one article, in which the authors stated that they conducted an experiment but without providing details regarding the nature of the experiment (i.e., being integrated into a survey or not).

Constructs and Definitions (RQ1)

Turning to the scales the articles used, we identified a total of 259 scales: 125 scales could be categorized as measuring source credibility, 74 scales as measuring message credibility, and 60 scales as measuring media credibility (cf. Table A3 in the Appendix).

While interest in media credibility has decreased since the start of credibility research, the prominence of source and message credibility has increased since the 1980s and has remained relatively constant since the turn of the century (Figure 2).
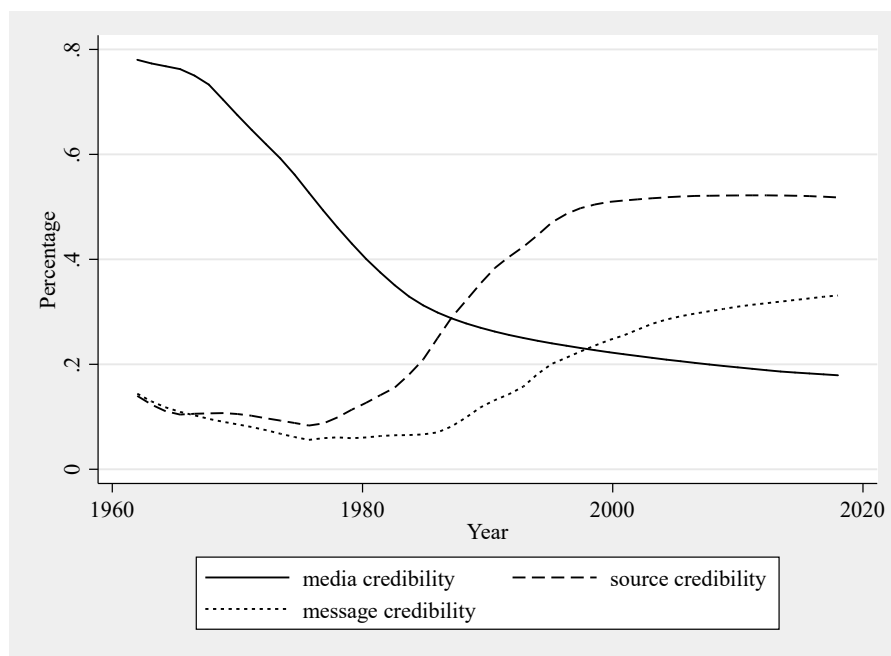
Figure 2. Constructs' relevance over time.

Scales of media credibility (n = 60) often compared different media channels (n = 27) or examined the credibility of a particular channel (e.g., the Web; n = 33). The most investigated media were newspapers (n = 42), TV (n = 40), and the Internet (n = 34). The relevance of source credibility increased steadily during the 1980s and1990s and seems to have remained stable since 2000 (n = 125), with studies focusing especially on collective sources (e.g., media organizations) since the 1990s (Figure A1 in the Appendix; n = 52). This may explain the decline in scholars' interest in media credibility: In the aftermath of the Roper (1985) studies, articles focused more on the credibility of media organizations than that of different media channels. Finally, most of the scales measuring source credibility focused on sources that the receiver did not know (e.g., no prominent people, friends, or other sources receivers would have had prior opinions about; n = 70). Message credibility only gained attention around the 1990s but has continuously been on the rise ever since. This may be related to the increased relevance of online communication: In the absence

of source information, the credibility of the information itself becomes increasingly important (Metzger et al., 2003). Scales investigating message credibility mostly focused on messages in the form of texts or nonmoving pictures (n = 61) and less on visual (moving; n = 10) or audio messages (n = 2). Only one contribution compared different types of messages (i.e., text, picture, and video).

Most of the scales identified across all constructs focused more on the online context (n = 147) and less on offline media, such as TV, radio, or newspapers (n = 70). Some scales focused on both the online and offline domains (n = 34). Over time, both the relevance of the offline context and the comparison of the two contexts decreased (Figure A2 in the Appendix), which is not very surprising given online communications' dominant role today.

Turning to credibility definition, only 79 of the 181 articles contained a definition (Table A3 in the Appendix) of the measured construct. Articles focusing on the online context more frequently included a definition (53.9%) than articles dealing with an offline context (32.6%).[9] However, most of the 79 definitions we identified were rather short and often focused on defining credibility in general instead of specifying the construct under investigation. Only 34 definitions contained a reference to the investigated construct. This reference was very generic in 14 of these 34 cases (e.g., scholars investigating a website's credibility referred to source credibility and not website credibility). A possible reason for only providing generic definitions was that many studies did not seek to explain credibility as comprehensively as possible but only focused on the effect that specific independent variables exert on credibility. Nevertheless, we identified five repeating patterns in credibility definitions: First, scholars considered credibility to be a perceptual variable and not an inherent quality of a given object (e.g., Lin, Spence, & Lachlan, 2016). Second, definitions frequently described credibility as multidimensional without providing further details (e.g., Magee & Kalyanaraman, 2010). Third, definitions highlighted that credibility could be ascribed to different objects such as sources or messages (e.g., Bracken, 2006). Fourth, scholars sometimes used the terms believability and credibility interchangeably (e.g., Lin & Spence, 2018). Depending on the construct of interest (i.e., source, media, or message), scholars used other terms than believability as synonyms for credibility. For message credibility, for example, some authors described credibility as the degree of trust in the accuracy of some information (Thon & Jucks, 2017) or as a judgment of some content's veracity (Appelman & Sundar, 2016). Finally, some authors also defined credibility as a situational judgment that was not enduring (e.g., Go et al., 2016; Kim, 2015).

While most articles did not discuss dimensionality from a theoretical perspective (whether the credibility construct consisted of one or more dimensions), many articles discussed dimensionality when they operationalized credibility (e.g., by using a scale for expertise and a scale for trustworthiness to measure source credibility). In total, 19.3% (n = 50) of all measured constructs were considered multidimensional. Source credibility was conceptualized as multidimensional (n = 43) most often, with scholars adopting either two (n = 17) or three (n = 25) dimensions. The frequency of multidimensional definitions was significantly higher for individual sources (n = 28) than for collective sources (n = 12; p = .002). Moreover, while scholars tended to agree that the primary dimensions of source credibility were

---

[9] Articles on offline communication are significantly less likely to contain a definition than articles focusing on online communication (p < .087).

trustworthiness (n = 33) and expertise (n = 23)/competence (n = 16), there was little agreement on the third dimension. The most frequently identified third dimensions included goodwill (n = 11), attractiveness (n = 9), and character (n = 6). Message (n = 2) and media (n = 5) credibility were only rarely considered multidimensional concepts.

Measurement Scales (RQ2)

As far as measurement scales were concerned, 130 articles used one scale, while 51 articles contained multiple scales. The latter type of articles measured different constructs (e.g., source and message credibility) with the same sample (i.e., parallel measurement), used different measures for the same construct, used different measures for different constructs with one or two samples, or used the same measure with different samples.

The majority (n = 156) of the 259 scales were based on existing scales but were so heavily adapted that one cannot speak of a replication (at least one-third of the original scale was adapted). In addition, 72 scales were replications (the same items are used) and 27 scales were adapted replications (less than one-third of the original scale was adapted). The majority (n = 80) of these 99 scales were direct replications, that is, no translation was necessary.[10] We could not determine the origin of four scales. (Adapted) Replications were most frequently used to measure source credibility (34.4% of scales). In general, many author(s) built on the works of Flanagin and Metzger (2000, 2007), Gaziano and McGrath (1986), McCroskey (1966), McCroskey and Teven (1999), McCroskey and Young (1981), and Meyer (1988).

The number of items used ranged from single-item measures (n = 34) up to measures based on 35 items (n = 1). Scales of source credibility generally consisted of more items (M = 7.9, SD = 6.1) than of media (M = 4.6, SD = 3.8) or message credibility scales (M = 5.0, SD = 3.5). This pattern is in keeping with source credibility's greater likelihood of being conceptualized as multidimensional compared with the other two constructs. Most measures were either additive indexes (n = 128) or semantic differential scales (n = 85), with the latter most often used to measure source credibility (n = 57). Most constructs were measured on 5- or 7-point scales (n = 183).

In general, measurements varied widely: We identified 198 different items in the 259 scales (Table A4 in the Appendix), 123 of which were only used once or twice. However, some of those items can be considered synonyms (e.g., objective and neutral or biased and slanted). Moreover, the high number of items can also be attributed to different translations into English. Overall, we identified several patterns in the items used.

Some items were used considerably more often than others. The items used most frequently (n ≥ 20) included trustworthy (n = 123), accurate (n = 97), believable (n = 78), credible (n = 71), fair (n = 54), unbiased (n = 50), expert (n = 49), honest (n = 44), reliable (n = 27), tells the whole story (n = 26), competent (n = 25), informed (n = 22), trained (n = 21), intelligent (n = 21), and complete (n = 21). At least a quarter of all message credibility scales contained the items accurate, believable, credible, fair,

---

[10] Replications might use answer scales different from the measures they replicate.

trustworthy, and unbiased. One-quarter of all media credibility scales consisted of the items accurate, believable, credible, fair, and trustworthy. Finally, at least 25% of all source credibility scales used the items credible, expert, honest, and trustworthy. The outline shows that message and media credibility overlapped on these items. This may be a result of message credibility scholars' reliance on Gaziano and McGrath's (1986) and Meyer's (1988) media credibility scales. However, using the same items for message and media credibility can be reasonable: Asking respondents to evaluate the believability of a message or a medium seems logical if the construct being evaluated is clear to the respondents.

We also found evidence for congruence between construct and measurement. First, some items appeared in scales for some constructs more often than for others: Scholars used "accurate," "complete," "tells the whole story," and "unbiased" to measure message credibility more often than to measure media or source credibility. "Believable" and "fair" appeared in message or media credibility scales more frequently than they appeared in source credibility scales. Finally, scholars typically used "competent," "expert," "honest," "informed," "intelligent," "reliable," "trained," and "trustworthy" for source credibility (Table A5 in the Appendix). Second, some items (i.e., "believable," "competent," "complete," "expert," "informed," "intelligent," "reliable," and "trained") appeared in online scales more frequently (in relative terms), whereas others (i.e., "fair," "tells the whole story," and "unbiased") were typical of offline scales (Table A6 in the Appendix). Hence, these results indicate that scholars measured credibility differently depending on the construct and the context.

On a more critical note, we observed that some items' conceptual link to credibility was not evident. For example, items like the "story" included major facts, or site organization predicted credibility more than they formed part of the construct itself. Moreover, some items like "sexy" or "sociable" seemed to describe more general perceptions rather than credibility. Finally, several items' exact meaning and relation to credibility remained unclear (e.g., website architecture).

To refine these findings, we conducted a Ward's (1963) linkage cluster analysis, which would allow us to identify scales that clustered together (RQ2). We preprocessed the data by merging synonymous items when they clearly referred either to source/media or to message.[11] We relied on the Oxford dictionary to define the synonyms (Oxford University Press, n.d.).[12] We used the Duda-Hart (Duda & Hart, 1973) criterion, the dendrogram, together with the interpretability of the cluster solutions to determine the appropriate number of clusters.

The cluster analysis suggests a three-cluster solution (Figure A3 in the Appendix). However, most scales clustered into the first (n = 156) and the third group (n = 68). Whereas the first group mainly contained source scales (48.1%), message credibility scales constituted the biggest group in cluster three (38.2%; Table 2). In fact, more than a third of all media and message credibility scales clustered in the third group (Table 3). The second group contained few scales (n = 17) and almost all scales in this group measured individual online source credibility (e.g., DeGroot, Young, & VanSlette, 2015; Lin & Spence, 2018; Lin et al., 2016; Yilmaz & Quintero Johnson, 2016). In addition, scales in this second cluster were largely based on the study by McCroskey and Teven (1999).

---

[11] Additional cluster analyses (e.g., without preprocessing) yielded largely similar results (see Appendix).

[12] A list of the regrouped items can be found in Table A7 in the Appendix.

As such, an interpretation along the three constructs (i.e., source, media, and message) is suitable to some extent. For example, individual sources were hardly ever measured using a cluster three scale (n = 4) but instead relied on scales from clusters one (n = 32) or two (n = 10). Moreover, the items most frequently used in cluster two scales all related to sources rather than to message (e.g., "caring," "intelligent"; Table A8 in the Appendix). Cluster three scales also had a certain consistency regarding the most frequently used items: These items referred to messages rather than to sources (e.g., "accurate" or "complete"). The scales in this cluster were often inspired by Gaziano and McGrath (1986) or Meyer (1988), who developed measures for media credibility. Nevertheless, cluster three scales also frequently relied on message-specific measures as provided by Flanagin and Metzger (2007).

Table 2. Cluster Attribution by Construct (Regrouped Items, Row Percent).

	Source N (%)	Media N (%)	Message N (%)	Total N (%)
Cluster 1	75 (48.1)	38 (24.4)	43 (27.6)	156 (100.0)
Cluster 2	17 (100.00)	0 (0.0)	0 (0.0)	17 (100.0)
Cluster 3	21 (30.9)	21 (30.9)	26 (38.2)	68 (100.0)

Table 3. Cluster Attribution by Construct (Regrouped Items, Column Percent).

	Source N (%)	Media N (%)	Message N (%)
Cluster 1	75 (66.4)	38 (64.4)	43 (62.3)
Cluster 2	17 (15.0)	0 (0.0)	0 (0.0)
Cluster 3	21 (18.6)	21 (35.6)	26 (37.7)
Total	113 (100.0)	59 (100.0)	69 (100.0)

Finally, we can also interpret the clusters along the communication context: Although each cluster mainly contained online scales, almost a third of the scales in clusters one and three measured credibility offline (Table 4). In contrast, save for two exceptions, cluster two scales all measured credibility online.

Table 4. Cluster Attribution by Context (Regrouped Items, Column Percent).

	Online N (%)	Offline N (%)	Other/Various N (%)	Total N (%)
Cluster 1	82 (55.4)	41 (27.7)	25 (16.9)	148 (100.0)
Cluster 2	15 (88.2)	2 (11.8)	0 (0.0)	17 (100.0)
Cluster 3	40 (58.8)	19 (27.9)	9 (13.2)	68 (100.0)

To sum up, the analysis of the scales reveals great heterogeneity in the ways in which credibility was measured. Moreover, the cluster analysis does not result in three clusters that mirror the three traditional constructs of source, media, and message credibility. Cluster three, for example, confirms that there is an overlap between media and message credibility. Moreover, cluster one indicates that the same items are used to measure source, message, and media credibility. Nevertheless, the cluster analysis confirms the frequency analysis in that patterns across the scales are in keeping with constructs and communication contexts. This is a sign of consistency in the measurement of credibility, but it also reveals a new potential dividing line between online and offline constructs—at least for the second cluster.

Quality of Measurement Scales

Based on the coded information, we also retrieved information about the quality of the measurement scales (i.e., reliability and validity).

One indicator of a measure's quality is Cronbach's α, which measures scales' internal consistency. Of all the studies reporting reliability scores (n = 192), 189 achieved good α values (i.e., > .70; min = .58, max = .98, M = .87, SD = .06). While extremely high α values can indicate redundancies in a scale (Streiner, 2003), only 15 scales achieved α values equal to or higher than .95. Alternatively, scholars have suggested reporting inter-item correlations along with Cronbach's α (Clark & Watson, 1995). However, these values are only provided for two scales. Still, the reported α values indicate that these scales have a relatively high internal consistency reliability.

Additional information about the quality of the measures can be derived based on whether researchers subjected their scales to validity tests (e.g., factor analysis) or pretested their scales before using them (e.g., to address issues of face validity). Such procedures are most important in the case of new or adapted scales. As we outlined earlier, we classified most of the identified scales as new scales because, despite relying on existing scales, they have been heavily adapted. Some of the author(s) using such new or adapted scales (n = 183) conducted validity tests (e.g., factor analysis; 28.4%) and a handful conducted pretests (8.2%). A majority, however, either did not perform or report on such tests.

We also evaluated the quality of the scales based on their external validity, which describes the scales' applicability beyond the study context. The relevant indicators are the type of the sample and the sample size. Scales applied to large samples (n ≥ 1,000) were rare. In fact, half of all scales were applied to samples smaller than or equal to 262 (min = 20, max = 6,738, M = 562.3, SD = 967.8). Moreover, scholars often relied on convenience samples. Hence, we know little about the degree to which these findings were generalizable to broader populations or populations beyond each study's context.

Finally, a detailed construct definition linked to its measurement is an important indicator that allows us to evaluate whether a measure adequately measures what it is supposed to measure (i.e., content validity; Carpenter, 2018). Nevertheless, authors did not provide construct-specific definitions for most of the identified scales. This omission makes it difficult to evaluate the coherence between the theoretical construct and its measurement.

Discussion

This article has sought to provide an overview of the state-of-the-art definitions (RQ1) and measures (RQ2) of credibility in media research over the past seven decades. Our overview aims to serve scholars to make informed decisions about the use of the relevant concept(s) in their own work.

In terms of credibility definitions (RQ1), we find that most authors only provided generic definitions—if they provided a definition at all. However, they did differentiate between the constructs of source, media, and message credibility, which suggests that there is a conceptual difference between the

three constructs. Moreover, while construct-specific definitions were almost nonexistent in the theoretical sections of the analyzed articles, they became apparent in the choice of items included in the operationalization and measurement of the constructs. More specifically, even though source credibility was the only construct regularly conceptualized as multidimensional, scholars did not agree on whether the construct was two- or three-dimensional. Many discussions pointed to a potential third dimension, in addition to the two core dimensions of expertise/competence and trustworthiness. However, the lack of construct-specific definitions and measurements that reflect such definitions may decrease scholars' incentives to replicate existing scales, as the absence of clear conceptualizations makes it difficult to decide that a particular scale is appropriate for one's purpose.

Indeed, our descriptive review shows that measurement varies widely (RQ2) and that replicating existing scales is relatively uncommon. Nevertheless, some clear patterns related to measurement do emerge. First, the credibility of individual sources in an online context is often measured based on McCroskey and Teven's (1999) scales (cluster 2), which suggests that the latter find wide application in an online environment. Second, message-specific measurements (cluster 3) are often based not only on scales developed to measure media credibility (e.g., Gaziano & McGrath, 1986; Meyer, 1988) but also on scales specifically aimed at capturing message credibility (e.g., Flanagin & Metzger, 2007). In general, however, the cluster analysis provides limited support for a consistent differentiation between source, message, and medium credibility: Scholars rely on the same items to measure collective source credibility, media credibility, and message credibility. Even though theoretical considerations would lead us to expect an overlap between source and media credibility we found that, surprisingly, message and source credibility seem to overlap in the first cluster as well. We also find an overlap in measurement scales for media and message credibility. While using the same items across the three constructs may seem reasonable in some cases, it may be more problematic in others. In any case, scholars should ensure that respondents know which construct they are asked to evaluate, especially when measuring various constructs simultaneously.

As far as the quality of the identified measures is concerned, we find that most scales reported high reliability scores. However, three factors complicated the evaluation of measurement quality. First, few studies reported results of validity tests or from scale pretesting. Second, most scales were used on small to moderate convenience samples. Third, most studies lacked construct-specific definitions linked to the measurement in use. Together, these factors made it difficult for researchers to assess the scales' content validity and applicability to other study contexts or broader populations.

Toward a New Model of Duality

Our analysis reveals a diversity of different measurement scales and an overlap in measurement along the three constructs. We do not maintain that credibility can and should be measured uniformly. On the contrary, credibility is an evaluation of an object or a subject in a specific situation (Go et al., 2016). As such, credibility perceptions are bound to the particular situation with its specific construct of evaluation. It, therefore, seems legitimate that the measurement of credibility is also adapted to that specific situation and the specific construct under investigation. However, the lack of clear conceptual links between concept and measurement makes it difficult to judge whether scholars' intention in scale construction was indeed to adapt it to the context and the construct at hand. It is also possible that the failure to establish a link

indicates that scholars rely on the trinity heuristic when they decide on a measurement. Consequently, the distinction between source, media, and message may constitute a temptation to take a shortcut toward operationalization without providing a link from concept to measurement.

We propose that we move toward a dual model and only distinguish between sources (i.e., personified/individual or objectified/collective) and messages (i.e., pieces of information/content). An anchor (i.e., personified) or a TV channel (i.e., objectified) would then be considered sources, while a newspaper article belongs to the message category. Consequently, we no longer need the third construct of media credibility. Our main argument for giving up the construct of media credibility is as follows: One group of media credibility scholars investigates the credibility of a medium in general (e.g., the "Internet"). The second group of media credibility scholars investigates specific media channels (e.g., *The New York Times*). In the former case, scholars investigate a general attitude and, therefore, trust rather than credibility. In the latter case, the object is, in our opinion, an objectified source more than a channel. Consequently, there is no conceptual need for a third construct.

Although our empirical analysis also shows an overlap between media and message credibility, integrating these two constructs is uncalled for from a conceptual point of view. As outlined, media credibility is often used to measure (objectified) sources such as specific newspapers, whereas message credibility measures the credibility of the content of a message. Believing a source or a message are two conceptually distinct acts: It is possible that one would consider a newspaper outlet in general as credible while rejecting a specific message transmitted via said outlet. Moreover, a message can be intrinsically believable even if coming from an untrustworthy source, for example, the truthfulness of the content is evidenced because of the act of communication itself (Sperber et al., 2010). Our proposition of a dual model also finds support from other authors such as Sperber and colleagues (2010), who distinguish between vigilance toward the source of communication (who to believe) and vigilance toward the content of the information (what to believe). Moreover, we believe that most scholars seem to already follow this duality in their measurements, without being explicit about it.

Depending on the category a credibility object belongs to, scholars should make sure they use corresponding items and clearly indicate to respondents the construct that is being evaluated. For instance, a credibility object in the "message" category would then be measured using items suitable for messages, such as "accurate" or "complete," whereas objects in the "sources" category would be measured with items suitable for either personified sources (e.g., anchor), such as "trained" or "intelligent," or for objectified sources (e.g., TV channel) such as "fair" or "believable." By this, we do not imply that there is a set of items that clearly belongs to any one category but that this choice should be conceptually driven and based on the chosen construct. Moreover, scholars should explicitly state these considerations.

Conclusion

Our review has two main limitations. First, we focused exclusively on articles published in English. Second, we included only communication journals. Conceptualizations and measures of credibility may differ between linguistic areas and research traditions. Despite these limitations, this review is to the best of our knowledge the first attempt to provide a comprehensive overview of conceptualizations and measures of

credibility. By doing so, it fills not only an important gap in the existing literature but also serves as a guide for future quantitative studies on credibility.

References

Appelman, A., & Sundar, S. S. (2016). Measuring message credibility: Construction and validation of an exclusive scale. *Journalism & Mass Communication Quarterly, 93*(1), 59–79. doi:10.1177/1077699015606057

Bentele, G., & Seidenglanz, R. (2008). Trust and credibility: Prerequisites for communication management. In A. Zerfass, B. van Ruler, & K. Sriramesh (Eds.), *Public relations research: European and international perspectives and innovations* (pp. 49–62). Wiesbaden, Germany: VS Verlag für Sozialwissenschaften.

Berdahl, L., Bourassa, M., Bell, S., & Fried, J. (2016). Exploring perceptions of credible science among policy stakeholder groups: Results of focus group discussions about nuclear energy. *Science Communication, 38*(3), 382–406. doi:10.1177/1075547016647175

Berlo, D. K., Lemert, J. B., & Mertz, R. J. (1969). Dimensions for evaluating the acceptability of message sources. *Public Opinion Quarterly, 33*(4), 563–576.

Borah, P. (2014). The hyperlinked world: A look at how the interactions of news frames and hyperlinks influence news credibility and willingness to seek information. *Journal of Computer-Mediated Communication, 19*(3), 576–590. doi:10.1111/jcc4.12060

Bracken, C. C. (2006). Perceived source credibility of local television news: The impact of television form and presence. *Journal of Broadcasting & Electronic Media, 50*(4), 723–741. doi:10.1207/s15506878jobem5004_9

Carpenter, S. (2018). Ten steps in scale development and reporting: A guide for researchers. *Communication Methods and Measures, 12*(1), 25–44. doi:10.1080/19312458.2017.1396583

Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment, 7*(3), 309–319. doi:10.1037/1040-3590.7.3.309

DeGroot, J. M., Young, V. J., & VanSlette, S. H. (2015). Twitter use and its effects on student perception of instructor credibility. *Communication Education, 64*(4), 419–437. doi:10.1080/03634523.2015.1014386

Duda, R. O., & Hart, P. E. (1973). *Pattern classification and scene analysis*. New York, NY: Wiley.

Ebscohost. (2018, October 31). *Communication abstracts: Database coverage list*. Retrieved from
        https://www.ebscohost.com/titleLists/cax-coverage.htm

Eisend, M. (2003). *Glaubwürdigkeit in der Marketingkommunikation. Konzeption, Einflussfaktoren und
        Wirkungspotenzial* [Credibility in marketing communication: Concept, influencing factors and
        impact potential]. Wiesbaden, Germany: Deutscher Universitätsverlag.

Fico, F., Richardson, J. D., & Edwards, S. M. (2004). Influence of story structure on perceived story bias
        and news organization credibility. *Mass Communication and Society, 7*(3), 301–318.
        doi:10.1207/s15327825mcs0703_3

Flanagin, A. J., & Metzger, M. J. (2000). Perceptions of Internet information credibility. *Journalism & Mass
        Communication Quarterly, 77*(3), 515–540. doi:10.1177/107769900007700304

Flanagin, A. J., & Metzger, M. J. (2007). The role of site features, user attributes, and information
        verification behaviors on the perceived credibility of web-based information. *New Media &
        Society, 9*(2), 319–342. doi:10.1177/1461444807075015

Freelon, D. G. (2010). ReCal: Intercoder reliability calculation as a web service. *International Journal of
        Internet Science, 5*(1), 20–33. Retrieved from
        https://dfreelon.org/publications/2010_ReCal_Intercoder_reliability_calculation_as_a_web_servic
        e.pdf

Fuoli, M., & Hart, C. (2018). Trust-building strategies in corporate discourse: An experimental study.
        Discourse & Society, 29(5), 514–552. doi:10.1177/0957926518770264

Gaziano, C., & McGrath, K. (1986). Measuring the concept of credibility. *Journalism Quarterly, 63*(3), 451–
        462. doi:10.1177/107769908606300301

Go, E., You, K. H., Jung, E., & Shim, H. (2016). Why do we use different types of websites and assign
        them different levels of credibility? Structural relations among users' motives, types of websites,
        information credibility, and trust in the press. *Computers in Human Behavior, 54*, 231–239.
        doi:10.1016/j.chb.2015.07.046

Golan, G. J. (2010). New perspectives on media credibility research. *American Behavioral Scientist, 54*(1),
        3–7. doi:10.1177/0002764210376307

Hong, T. (2006). The influence of structural and message features on web site credibility. *Journal of the
        American Society for Information Science and Technology, 57*(1), 114–127.
        doi:10.1002/asi.20258

Hovland, C. I., Janis, I. L., & Kelly, H. H. (1953). *Communication and persuasion: Psychological studies of
        opinion change*. New Haven, CT: Yale University Press.

Iyengar, S., & Valentino, N. A. (2000). Who says what? Source credibility as a mediator of campaign advertising. In A. Lupia & M. McCubbins (Eds.), *Elements of reason: Cognition, choice, and the bounds of rationality* (pp. 109–129). Cambridge, UK: Cambridge University Press.

Johnson, T. J., & Kaye, B. K. (2016). Some like it lots: The influence of interactivity and reliance on credibility. *Computers in Human Behavior, 61*, 136–145. doi:10.1016/j.chb.2016.03.012

Kim, M. (2015). Partisans and controversial news online: Comparing perceptions of bias and credibility in news content from blogs and mainstream media. *Mass Communication and Society, 18*(1), 17–36. doi:10.1080/15205436.2013.877486

Kiousis, S. (2001). Public trust or mistrust? Perceptions of media credibility in the information age. *Mass Communication and Society, 4*(4), 381–403. doi:10.1207/S15327825MCS0404_4

Kohring, M., & Matthes, J. (2007). Trust in news media: Development and validation of a multidimensional scale. *Communication Research, 34*(2), 231–252. doi:10.1177/0093650206298071

Lin, X., & Spence, P. R. (2018). Identity on social networks as a cue: Identity, retweets, and credibility. *Communication Studies, 69*(5), 461–482. doi:10.1080/10510974.2018.1489295

Lin, X., Spence, P. R., & Lachlan, K. A. (2016). Social media and credibility indicators: The effect of influence cues. *Computers in Human Behavior, 63*, 264–271. doi:10.1016/j.chb.2016.05.002

Lock, I., & Seele, P. (2017). Measuring credibility perceptions in CSR communication: A scale development to test readers' perceived credibility of CSR reports. *Management Communication Quarterly, 31*(4), 584–613. doi:10.1177/0893318917707592

Lupia, A. (2000). Who can persuade whom? How simple cues affect political attitudes. In J. H. Kuklinski (Ed.), *Thinking about political psychology* (pp. 51–88). Cambridge, UK: Cambridge University Press.

Magee, R. G., & Kalyanaraman, S. (2010). The perceived moral qualities of web sites: Implications for persuasion processes in human-computer interaction. *Ethics and Information Technology, 12*(2), 109–125. doi:10.1007/s10676-009-9210-1

McCroskey, J. C. (1966). Scales for the measurement of ethos. *Speech Monographs, 33*, 65–72. doi:10.1080/03637756609375482

McCroskey, J. C., & Teven, J. J. (1999). Goodwill: A reexamination of the construct and its measurement. *Communication Monographs, 66*(1), 90–103. doi:10.1080/03637759909376464

McCroskey, J. C., & Young, T. J. (1979). The use and abuse of factor analysis in communication research. *Human Communication Research, 5*(4), 375–382. doi:10.1111/j.1468-2958.1979.tb00651.x

McCroskey, J. C., & Young, T. J. (1981). Ethos and credibility: The construct and its measurement after three decades. *Central States Speech Journal, 32*(1), 24–34. doi:10.1080/10510978109368075

Metzger, M. J., Flanagin, A. J., Eyal, K., Lemus, D. R., & McCann, R. M. (2003). Credibility for the 21st century: Integrating perspectives on source, message, and media credibility in the contemporary media environment. *Annals of the International Communication Association, 27*(1), 293–335. doi:10.1080/23808985.2003.11679029

Meyer, H. K., Marchionni, D., & Thorson, E. (2010). The journalist behind the news: Credibility of straight, collaborative, opinionated, and blogged "news." *American Behavioral Scientist, 54*(2), 100–119. doi:10.1177/0002764210376313

Meyer, P. (1988). Defining and measuring credibility of newspapers: Developing an index. *Journalism & Mass Communication Quarterly, 65*(3), 567–574. doi:10.1177/107769908806500301

Newhagen, J., & Nass, C. (1989). Differential criteria for evaluating credibility of newspapers and TV news. *Journalism Quarterly, 66*(2), 277–284. doi:10.1177/107769908906600202

Ohanian, R. (1990). Construction and validation of a scale to measure celebrity endorsers' perceived expertise, trustworthiness, and attractiveness. *Journal of Advertising, 19*(3), 39–52. doi:10.1080/00913367.1990.10673191

Oxford University Press. (n.d.). *Oxford English Dictionary*. Retrieved from https://en.oxforddictionaries.com/thesaurus/

Paré, G., Trudel, M.-C., Jaana, M., & Kitsiou, S. (2015). Synthesizing information systems knowledge: A typology of literature reviews. *Information & Management, 52*(2), 183–199. doi:10.1016/j.im.2014.08.008

Peters, H. P. (1992). The credibility of information sources in West Germany after the Chernobyl disaster. *Public Understanding of Science, 1*(3), 325–343. doi:10.1088/0963-6625/1/3/006

Petty, R. E., & Cacioppo, J. T. (1981). *Attitudes and persuasion: Classic and contemporary approaches*. New York, NY: Routledge.

Pjesivac, I., Geidner, N., & Cameron, J. (2018). Social credibility online: The role of online comments in assessing news article credibility. *Newspaper Research Journal, 39*(1), 18–31. doi:10.1177/0739532918761065

Pjesivac, I., & Rui, R. (2014). Anonymous sources hurt credibility of news stories across cultures: A comparative experiment in America and China. *International Communication Gazette, 76*(8), 641–660. doi:10.1177/1748048514548534

Roper Organization. (1985). *America's watching: 30th anniversary 1959–1989*. New York, NY: Television Information Office.

Rosenthal, P. I. (1971). Specificity, verifiability, and message credibility. *Quarterly Journal of Speech, 57*(4), 393–401. doi:10.1080/00335637109383084

Rowley, J., Johnson, F., & Sbaffi, L. (2015). Students' trust judgements in online health information seeking. *Health Informatics Journal, 21*(4), 316–327. doi:10.1177/1460458214546772

Schweiger, W. (2000). Media credibility—Experience or image? A survey on the credibility of the world wide web in Germany in comparison to other media. *European Journal of Communication, 15*(1), 37–59. doi:10.1177/0267323100015001002

Seo, H., & Lim, J. (2010). A comparative study on source credibility and use in multinational nuclear talks. *Asian Journal of Communication, 20*(4), 440–455. doi:10.1080/01292986.2010.496861

Simons, T. (2002). Behavioral integrity: The perceived alignment between managers' words and deeds as a research focus. *Organization Science, 13*(1), 18–35. doi:10.1287/orsc.13.1.18.543

Smith, R. G. (1978). *The message measurement inventory: A profile for communication analysis*. Bloomington: Indiana University Press.

Sperber, D., Clément, F., Heintz, C., Mascaro, O., Mercier, H., Origgi, G., & Wilson, D. (2010). Epistemic vigilance. *Mind & Language, 25*(4), 359–393. doi:10.1111/j.1468-0017.2010.01394.x

Streiner, D. L. (2003). Starting at the beginning: An introduction to coefficient alpha and internal consistency. *Journal of Personality Assessment, 80*(1), 99–103. doi:10.1207/S15327752JPA8001_18

Sundar, S. S. (1999). Exploring receivers' criteria for perception of print and online news. *Journalism & Mass Communication Quarterly, 76*(2), 373–386. doi:10.1177/107769909907600213

Sylvester, A., Tate, M., & Johnstone, D. (2013). Beyond synthesis: Re-presenting heterogeneous research literature. *Behaviour & Information Technology, 32*(12), 1199–1215. doi:10.1080/0144929X.2011.624633

Thon, F. M., & Jucks, R. (2017). Believing in expertise: How authors' credentials and language use influence the credibility of online health information. *Health Communication, 32*(7), 828–836. doi:10.1080/10410236.2016.1172296

Trilling, D. (2019). Conceptualizing and measuring media exposure as network of users and news items. In C. Peter, T. Naab, & R. Kühne (Eds.), *Measuring media user and exposure. Recent developments and challenges* (pp. 297–317)*. Cologne, Germany: Herbert von Halem Verlag.

Tsfati, Y., & Cappella, J. N. (2003). Do people watch what they do not trust? Exploring the association between news media skepticism and exposure. *Communication Research, 30*(5), 504–529. doi:10.1177/0093650203253371

Wallander, L. (2009). 25 years of factorial surveys in sociology: A review. *Social Science Research, 38*(2009), 505–520. doi:10.1016/j.ssresearch.2009.03.004

Ward, J. H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, *58*(301), 236–244. doi:10.1080/01621459.1963.10500845

West, M. D. (1994). Validating a scale for the measurement of credibility: A covariance structure modelling approach. *Journalism Quarterly, 71*(1), 159–168. doi:10.1177/107769909407100115

Winter, S., & Krämer, N. C. (2014). A question of credibility: Effects of source cues and recommendations on information selection on news sites and blogs. *Communications, 39*(4), 435–456. doi:10.1515/commun-2014-0020

Yilmaz, G., & Quintero Johnson, J. M. (2016). Tweeting facts, Facebooking lives: The influence of language use and modality on online source credibility. *Communication Research Reports, 33*(2), 137–144. doi:10.1080/08824096.2016.1155047