A Little Bird Told Me, So I Didn’t Believe It: Twitter, Credibility, and Issue Perceptions

Mike Schmierbach & Anne Oeldorf-Hirsch

This article investigates how media use of the microblogging tool Twitter affects perceptions of the issue covered and the credibility of the information. In contrast to prior studies showing that ordinary blogs are often judged credible, especially by their users, data from 2 experiments show that Twitter is considered less credible than various forms of stories posted on a newspaper Web site, and fails to convey importance as well as a newspaper or blog.

Keywords: Credibility; Information Seeking; Twitter

In the last decade, individuals have increasingly turned to the Internet for information, including the news. From 1994 to 2008 reading the news online (at least 3 days per week) increased from 2% to 37% (Pew Research Center, 2008). In turn, news organizations have struggled to adjust both to competition from new outlets and to distribution through new technologies. Individuals get information directly from organization Web sites, but also indirectly through individual or third-party blogs that re-post news. Somewhere between these lies the microblogging tool Twitter, which both media outlets and others can use to provide quick tweets alerting followers to the day’s headlines and directing them to full stories.

In this context, judgments about credibility are both critical and complex. Faced with a diversity of news sources and stories, individuals must decide how much they trust information and whether to act on it. Yet, new technologies can mask or complicate the underlying source of information. Consider, for example, a news story that...
initially appears in *The New York Times*. A person might find the story on the organization’s Web site, find a link from a prominent blog, be sent an e-mail recommending the story or the blog post, or come upon a tweet referencing the story. In such circumstances, opinions about the initial news outlet may become intertwined with opinions about blogs or Twitter, as well as the individuals making the recommendation.

In this study, we consider how these complications may affect judgments about the credibility of information posted to Twitter, as well as further interest in the story topic. By comparing evaluations of Twitter posts to evaluations of the original story in different formats and locations, we explore how a specific microblogging technology influences perceptions of a news story.

Blogs, Microblogs, and News

Before considering the specific microblogging tool of Twitter, we first consider the key antecedent technology. Blogs have been defined as personal Web sites that are regularly updated and contain archived dated entries in reverse chronological order, which contain primarily text, but may also contain photos or other multimedia, and usually allow for audience comments (Herring, Scheidt, Bonus, & Wright, 2004; Herring, Scheidt, Wright, & Bonus, 2005; Nardi, Schiano, & Gumbrecht, 2004). Technorati, the first blog search engine and leading authority on blogging, has indexed 133 million blogs in 88 languages in at least 66 countries (see “State of the Blogosphere 2008” at http://technorati.com/blogging/feature/state-of-the-blogosphere-2008/). As of 2009, a reported 77% of active Internet users read blogs regularly, totaling up to 346 million blog readers (see “State of the Blogosphere 2009” at http://technorati.com/blogging/feature/state-of-the-blogosphere-2009/).

Only a small portion of these blogs are news focused; the most popular topic for blogs continues to be the author’s personal life, and few serve the “filtering” and link-sharing role often associated with blogs in the political communication domain (Wei, 2009). Similarly, most blogs have no professional aspirations, and the majority (72%) of bloggers consider themselves “hobbyists,” writing for personal satisfaction (see “State of the Blogosphere 2009” at http://technorati.com/blogging/feature/state-of-the-blogosphere-2009/). Readers have differing motivations, particularly information seeking and media checking, convenience, personal fulfillment, political surveillance, social surveillance, and expression and affiliation (Kaye, 2005). These motivations may explain why even personal blogs often have informational elements and why many of the most widely read blogs are more explicitly informative or news focused.

Nonetheless, blogs are not a primary channel for media sources to spread hard news, no doubt in part because they rely on the blog owner to pass on stories and, thus, provide context that the press cannot control. Instead, many news organizations have turned to microblogging and social media to share content. Microblogging is a hybrid of blogging and instant messaging in which updates form a feed, similar to a
blog, and users can quickly reply to or re-post others’ updates. The most popular service is Twitter, a site that allows users 140 characters to send tweets to their followers. Users can enter text and shortened URLs that lead to pictures or other sites using the Twitter homepage, their mobile phone, or one of hundreds of microblogging applications that show up on their desktop, their browser, or other Web sites (for a list of more than 260 applications that allow Twitter monitoring, see Goldstein, 2011). Other users can then choose to “follow” this user, meaning that they receive all of that person’s updates in an ongoing feed. Alternatively, they can also visit the Twitter site or another page where the individual feed can be displayed.

As of 2011, Twitter attracts at least 190 million users per month (Schonefeld, 2010), who post 150 million tweets per day (Siegler, 2011). Although daily chatter is common, news reporting and information sharing are other popular uses of this and similar microblogging platforms (Java, Song, Finin, & Tseng, 2007). Twitter’s identity as a news source is rapidly becoming more prominent, and the service has been recognized as a useful news and current events tool (Kwak, Lee, Park, & Moon, 2010; Phelan, McCarthy, & Smyth, 2009; Sakaki, Okazaki, & Matsuo, 2010). In fact, data collected and analyzed from millions of tweets on thousands of trending topics found that over 85% of trending topics on the site are headline news topics or persistent news topics (Kwak et al., 2010).

Considerable attention has been paid to traditional blogs and news use, and a growing amount of work considers Twitter as well. It is beyond the scope of this article to consider the full range of such research, but several studies highlight the comparisons offered between Twitter and more traditional journalism outlets. Much of this research focuses on the use of Twitter by so-called citizen journalists, demonstrating how the technology allows ordinary people the opportunity to shape and comment on the mainstream media agenda (Hermida, 2010), although some question its utility in the face of low levels of adoption (Murthy, 2011). Since its creation, Twitter emerged as a user-generated source of real-time news, often breaking news before any other media source, as in the Iran protests in 2009 (Grossman, 2009). Ampofo, Anstead, and O’Loughlin (2011) argued that debate viewers used Twitter to directly comment on and shape the dialogue about political discourse.

At the same time, research suggests that journalists are also increasingly enthusiastic about Twitter (Arceneaux & Weiss, 2010) and are willing to make use of the technology in promotion (Butcher, 2009; Oberholtzer, 2011). Ahmad (2010) argued that Twitter would function as a useful promotional tool for journalists, but the approach was driven by critical theory rather than empirical evidence. A content analysis by Muralidharan, Rasmussen, Patterson, and Shin (2011) concluded that media outlets do use Twitter and other social media to disseminate content, but often fail to take advantage of the potential for interaction with the audience. Perhaps the most notable studies in this area were conducted by C. F. Greer and Ferguson (2011; see also Ferguson & Greer, 2011), which evaluate the use of Twitter by television and radio stations, respectively. For example, in their study of local television use of the technology, these authors found that, by far, the most common use was in passing on news, rather than other promotional techniques. Although such studies can catalog
how media outlets have employed Twitter, they tell us little about how audiences respond. As Chen (2011) pointed out, social gratifications are a key motivator for Twitter use, and one-directional news distribution may be unlikely to serve such motivations. Thus, questions about the value of Twitter for journalists remain, and at present the most common journalistic use of Twitter appears to be as a means to promote or disseminate content rather than engage in social conversations, suggesting traditional news outlets may not be realizing the full potential of social media.

Finally, a study by Schultz, Utz, and Göritz (2011) presented a design most similar to the approach taken in this study, comparing the effects of different message strategies on audiences. In this study, participants were presented with crisis communication from a company through a newspaper, a blog or Twitter. Although the results were varied, in general individuals receiving a message via Twitter were less likely to share it with others and had a lower impression of the company. Thus, this research at least hints that the promise of Twitter might also hold some peril. We consider this in the context of news media and the effects Twitter could have on the perceived credibility of the outlet and its messages.

Credibility Assessments

Credibility is generally defined as the extent to which individuals find information or its source believable, accurate, and trustworthy (Flanagin & Metzger, 2000). Much of the prior research exploring the perceived credibility of Internet-based information has shown that online information is considered as or more credible than offline content. For example, Johnson, Kaye, Bichard, and Wong (2008) found that some online news sources were rated more credible than their offline counterparts. Kioussis (2001) found that online news was rated more credible than television, although less credible than newspapers. Flanagin and Metzger (2000) reached a similar conclusion. Thus, it appears that online information can be credible. As an Internet-based tool, Twitter may be evaluated similarly to other online information sources. In particular, as a type of microblogging, Twitter may be perceived in a manner consistent with other blogs. Research on blog credibility provides a mixed picture. A series of studies by Johnson and colleagues (Johnson & Kaye, 2004, 2009; Johnson et al., 2008) indicated that blogs were widely considered credible, even by comparison to more traditional news sources. However, these judgments were largely held by regular blog users; non-users had lower opinions of blog credibility. In addition, other research suggests that traditional online news sources sometimes rank higher than other online sources in terms of credibility (Flanagin & Metzger, 2007; Melican & Dixon, 2008).

One complicating factor for both Twitter and traditional blogs is that the source of information can be unclear. As Sundar (1999) pointed out, whereas traditional measures of credibility are mostly applied to content producers, in an online environment people assess individual pieces of information on the basis of various cues about credibility. These include not only the original source of information but also what Sundar and Nass (2001) called the selecting source—that is, the individual or
organization responsible for suggesting a story to a reader but not for gathering the
information initially. Their study found more favorable evaluations for information
selected by other users, rather than expert editors or oneself. This could help explain
the favorable ratings blogs receive in some studies. In a study of health messages, Hu
and Sundar (2010) considered both the originating source (e.g., a doctor or layper-
son) and the selecting source (e.g., a specific Web site, blog, or personal page). Their
study found no direct effect of either. Given that other research (Eastin, 2001) has
found that health information from experts is seen as more credible, it could be that
the selecting aspect of the process acted to partly mask normal source effects. How-
ever, Sundar, Knobloch-Westerwick, and Hastall (2007) found that the originating
source still influenced credibility when individuals were presented with information
on a news-aggregating site.

Compared with blogs and news aggregation services, Twitter represents a more
complex, ambiguous case. A traditional blog that posts links to news content clearly
has both an originating source and a selecting source. However, when news outlets
use Twitter to promote their content, these lines blur. For example, the official news
feed of The New York Times does not qualify as a “selecting source” as technically
defined in prior research because no third party is making judgments about story
appropriateness or importance. Yet, it could be that individuals will nonetheless
allow their feelings about Twitter to influence their assessment of the credibility of
the information source, although Twitter, as an entity, did not actually select the
tweeted content. If so, this would suggest that individuals do not necessarily make
credibility judgments through a careful consideration of sources and message content
but, rather, allow superficial cues to sway their judgments. If so, these cues may also
carry over to judgments about message credibility, if readers do not attend carefully to
content features in forming such judgments. On the other hand, some individuals
appear to distinguish between message and source credibility (J. D. Greer, 2003;
Kiousis, 2006). Kiousis measured message and source credibility of online news sep-
arately and found that the use of a Web site’s multimedia features had a significant,
positive influence on the perceived credibility of the source, but not the message. This
illustrates that even when a story’s credibility is seemingly unaffected by the medium,
perceptions of the source may be influenced. In a study on advertising’s effects on an
online story’s credibility, J. D. Greer found that, although message and source credi-
bility perceptions are correlated, the extent to which they are related differs by story
context. Thus, Twitter could affect judgments about source credibility to the extent
that it is perceived as a selecting or even originating source. In turn, these evaluations
might carry over to feelings about the message, but the literature offers reasons for
cautious and distinct assessment of both variables.

Hypotheses

Predicting the effects of Twitter requires combining seemingly disparate information
from several lines of research on credibility. Some of this work shows Internet sources
can be seen as quite credible. Even blogs, which are similar to Twitter in important ways, are often rated as trustworthy, although primarily by users or those otherwise disinclined to solely rely on mainstream media (Choi, Watt, & Lynch, 2006). However, blogs present “user-selected” content whereas Twitter may not, at least in the case of a feed from the media outlet itself. The literature also shows that people may evaluate the message and the source in distinct ways, and that the originating source (in this study, The New York Times) may be less influential than opinions about the selecting source.

Thus, evaluations of Twitter specifically may be important in shaping perceptions of both source and message; and, in this regard, Twitter may have factors working against it. As noted earlier, Shultz et al. (2011) found individuals exposed to corporate crisis communication via Twitter had somewhat more negative impressions of the company, suggesting that people may not respond as favorably to communication via that medium. A study by Dutta-Bergman (2004) found that jargon-dependent or incomplete health messages were seen as less credible. Given the brevity of Twitter messages even compared with traditional blogs, this might suggest the typical reader will respond negatively to typical tweets compared with complete news stories. Twitter credibility may also suffer because of perceptions of its use by and for celebrities, given their strong presence and “most-followed” status on the site (Hargittai & Litt, 2011; Marwick & Boyd, 2011; Wu, Hofman, Mason, & Watts, 2011). To evaluate the specific effects of “tweeting” on credibility judgments, we compare responses to a story located on a newspaper Web site to judgments about a tweet from that same newspaper advertising the story. Although the true source remains constant, it may be that individuals will perceive Twitter as a kind of quasi-selecting source, affecting credibility judgments. The message format could also influence these judgments. Because Twitter provides fewer markers of being a traditional news source and gives only a brief synopsis of a story, we expect that, unlike traditional blogs, Twitter posts will not be viewed favorably, either in terms of the message itself or the overall source credibility, as reflected in the following hypotheses:

**H1a:** Individuals will perceive the message as less credible when presented as a Twitter post by a news organization than as a story posted on the organization’s Web site.

**H1b:** Individuals will perceive the source as less credible when presented as a Twitter post by a news organization than as a story posted on the organization’s Web site.

By itself, credibility is a noteworthy outcome, and it is the primary focus of this study. Credibility is often cited as a factor linked to declining media use (Johnson et al., 2008), and it is a rich variable worthy of investigation. However, declines in credibility may also undercut the ability of media to shape issue perceptions even among those who continue to consume news (Wanta & Hu, 1994). If the Twitter posting reduces credibility, it could also reduce perceived importance. In addition, by including less information, the Twitter post and the shorter story would also...
rob individuals of the context needed to determine that a story was important. Story length could also be a heuristic showing the importance of the issue—as demonstrated in Graber’s (1988) study showing that individuals use length as one factor in deciding which news stories matter. Overall, this leads to the following hypothesis:

**H2:** Individuals will perceive content as more important when presented as a longer story than either as a short story or a Twitter post by a news organization.

Ultimately, one of the reasons news organizations use Twitter is to try to increase readership. The hope, presumably, is that potential readers will be intrigued by the brief headline and click the accompanying link to learn more, pushing up the number of visitors to the organization’s site. Certainly, this seems like a logical prospect. By limiting information, such brief tweets might increase uncertainty. Uncertainty Reduction Theory (Berger & Calabrese, 1975; Sept, Hildebrand, & Wexler, 1992) proposes that individuals will seek to resolve the anxiety and discomfort caused by this uncertainty by seeking information, and some research on media suggests at least limited circumstances in which this might occur (Boyle et al., 2004; Kubey & Peluso, 1990). On the other hand, if credibility or importance is seen as lower for the Twitter condition, this would probably make people less interested in pursuing the story further, particularly through the same news source. Ultimately, the current literature simply is not sufficient to make a clear prediction, but potential changes in information seeking as a result of Twitter exposure are important, as we explore in the following research question:

**RQ1:** How will intended information seeking after exposure to a Twitter post differ compared with exposure to a version of the story on the news organization Web site?

### Study 1

**Method**

Participants were recruited by students at a large, Northeastern U.S. university. Each student was required to recruit at least 15 volunteers. These individuals were contacted by e-mail and directed to an online consent form. Consenting participants were randomly assigned to one of three story conditions. After dropping a small number of participants who incorrectly answered a question about the story topic, the total *N* was 225.

Most of these participants were current students (71.1%), primarily at the same school as those recruiting (54.7%). The majority of the remaining participants had completed college (20.9%). Participants were predominantly women (63.6%) and White (88.9%). The participants were generally young (*M* = 25.36, *SD* = 11.02), but ages ranged from 18 to 74. Only 28.4% of participants used Twitter. Although this is a low rate, it is substantially higher than overall reported usage; according to research by the Pew Internet and American Life Project (Smith & Rainie, 2010),
just 8% of Internet users and 14% of those under 30 reported using Twitter when surveyed in November 2010, and rates were most likely lower during the time of data collection in the Fall of 2009.

As noted, participants viewed one of three versions of a news story adapted from *The New York Times*. Participants were presented with a screen-captured image of the story, which described a speech by President Obama on health care. The long form of the story provided the first four paragraphs of this story as they appeared on the “Money & Policy” section of the newspaper’s Web site. All ads were removed, as was material indicating the date. A “next page” box appeared at the bottom of the story, which had the headline, “Seeking to Woo the Insured, Obama Cites Risk of Losing Coverage.” In contrast to this version of the story, the short version included just the first paragraph, and the box at the bottom of the story said “read story.” Otherwise, the format matched. Finally, the Twitter version showed the “nytimes” Twitter feed page on Twitter’s site, and listed the tweet for this story: “Obama Keeps Up Health Care Push, Citing Uninsured,” along with a bit.ly link and the note that the tweet was made recently. Although hosted by Twitter, this page clearly indicates it is associated with *The New York Times*. For all three versions, the story or tweet was the actual text posted by the newspaper, with no alterations other than shortening the material. For all versions, other tweets, headlines, and stories were deleted.

For all three versions, participants were told the image was a screen capture of a recent news item that appeared online, and that they should look at it before clicking to proceed. In total, 52 individuals read the Twitter version, 77 read the short version, and 96 saw the long version. Participants were randomly assigned to conditions. Although true randomization means small differences in cell sizes could occur by chance, the inequality in numbers between conditions suggests higher numbers of participants in the short-story and Twitter versions dropped out prior to submitting their responses. Because they did not complete the study, we cannot be sure how they may have differed, but informal feedback from some participants suggests a portion of those in the Twitter and short-story versions may have believed the displayed story was incomplete or broken. If anything, this would suggest our results may underestimate negative feelings toward this condition, but we cannot be certain, and potential inequalities between conditions are a concern. To help address this, several control variables are included in analysis, including age, student status, prior use of Twitter, and gender, as described earlier.

**Dependent Variables**

*Source credibility.* Participants were asked to indicate how credible they found the source of the information. “Source” was used as an intentionally ambiguous label, allowing participants to perceive source in their own terms, rather than forcing them to evaluate either *The New York Times* by itself or explicitly prompting an assessment that included Twitter. This also ensured equivalent wording across conditions. An index was formed based on two questions, which were presented on a 7-point scale ranging from 1 (strongly disagree) to 7 (strongly agree) (\(M = 5.30, \ SD = 1.42;\)
Cronbach’s $\alpha = .91$). Specifically, participants were asked how much they agreed that they “trust the source of this information” and “believe this source is credible.”

**Message credibility.** Participants were also asked to indicate how credible they found the specific message they were presented with, using an index based on two questions presented on the same 7-point scale ($M = 5.04, SD = 1.43; \alpha = .93$). Specifically, participants were asked how much they agreed that they “believe the content of this text is true” and “trust the specific information in this text.”

**Issue importance.** Participants were asked how important they found health care reform, which likely indicates both involvement in the topic and support for reform. An index was formed based on six questions, all of which were presented on the 7-point agreement scale ($M = 5.39, SD = 1.35; \alpha = .93$). Specifically, participants were asked how much they agree that health care reform and (separately) the number of uninsured is important to them personally, that it is important to the nation as a whole, and that it will be important in 10 years.

**Information-seeking intent.** Participants were asked how likely they were to pursue this issue further. An index was formed based on nine questions, all of which were presented on a 7-point scale ranging from 1 (not likely) to 7 (very likely) ($M = 4.38, SD = 1.34; \alpha = .89$). Participants indicated how likely they were to take the following actions regarding “the issue of health care and the number of uninsured”:

- Click the link found in the story
- Read additional online stories
- Search for information
- Read additional stories in a print newspaper
- Watch network television news coverage
- Watch cable news coverage
- Read blogs
- Talk with friends
- Talk with family

**Information satisfaction.** Finally, participants were asked how “satisfied [they were] with the information found in the text.” This was measured with a single question, as described earlier, using a 7-point agreement scale ($M = 3.64, SD = 1.64$).

**Study 1 Results**

As noted, analysis included four control variables: age, prior use of Twitter, gender, and student status. All hypotheses were initially tested using an analysis of covariance (ANCOVA), with these control variables as covariates and story condition as the independent variable. $H_{1a}$ predicted that people would see stories posted on *The New York Times* Web site as more credible than those posted via Twitter. The analysis shows no significant effect of story condition on message credibility, $F(2, 184) = 1.13$, *ns.* As shown in Table 1, perceived message credibility, although somewhat lower for those in the Twitter condition, did not significantly vary between groups. Thus, $H_{1a}$ was not supported. $H_{1b}$ predicted that people would see stories posted on *The New York Times* Web site as having a more credible source than those posted via Twitter, although they also come from *The New York Times*. This finding was supported. The analysis showed an overall significant effect
of story condition on source credibility, $F(2, 184) = 4.90, p < .01 (\eta^2_p = .05)$. As shown in Table 1, perceived source credibility was significantly lower in the Twitter condition than in either the short- or long-story conditions, which did not significantly differ from one another.

H2 predicted that individuals would see the issue as more important if presented with a long story. This hypothesis was not supported, as the analyses found no significant difference between story conditions, $F(2, 187) = 0.24, ns$. Finally, RQ1 asked how story length and location would affect information-seeking intention. The analyses show no evidence that length or location influenced intended information seeking, $F(2, 186) = 0.31, ns$. The literature suggested potentially competing mechanisms by which length and location would affect information seeking. One possible mechanism, importance, was not influenced. Additional analysis confirms that condition had an effect on the alternative mechanism of information satisfaction, $F(2, 184) = 6.54, p < .01 (\eta^2_p = .07)$. However, as Table 1 shows, the levels of satisfaction did not match expectations. Whereas those who saw the short story were noticeably less satisfied than those reading the long story, those who saw only Twitter were as satisfied as those reading the long story. Stories posted on Twitter are seen as less credible, but fail to evoke curiosity about the issue.

To further explore the results, a regression analysis was run using issue importance, credibility, information satisfaction, and the control variables to predict information seeking. Perceived importance was a positive predictor of intent to seek future information ($\beta = .36, p < .001$), whereas neither source ($\beta = -.46, ns$) nor message credibility ($\beta = .17, ns$) were significant predictors. Information satisfaction was unrelated to information seeking ($\beta = .01, ns$). It is possible the null findings for credibility are a reflection of colinearity, as the two measures are strongly correlated with one another ($r = .85, p < .001$). Thus, we evaluated the zero-order correlation of both measures of credibility to information seeking, and found that both source credibility ($r = .38, p < .001$) and message credibility ($r = .38, p < .001$) were significantly related to information-seeking intention. Thus, it may be that story location and length had an indirect effect on information seeking by way of credibility judgments, but the evidence for this is weak.

### Table 1  Credibility, Importance, Information-Seeking, and Information Satisfaction Scores by Post Length and Location (Study 1)

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>Twitter</th>
<th>Short story</th>
<th>Long story</th>
</tr>
</thead>
<tbody>
<tr>
<td>Message credibility</td>
<td>4.77</td>
<td>5.11</td>
<td>5.18</td>
</tr>
<tr>
<td>Source credibility</td>
<td>4.76</td>
<td>5.47</td>
<td>5.56</td>
</tr>
<tr>
<td>Importance</td>
<td>5.43</td>
<td>5.36</td>
<td>5.51</td>
</tr>
<tr>
<td>Information seeking</td>
<td>4.41</td>
<td>4.38</td>
<td>4.35</td>
</tr>
<tr>
<td>Information satisfaction</td>
<td>3.92</td>
<td>3.04</td>
<td>3.92</td>
</tr>
</tbody>
</table>

Note. Scores shown are estimated means with age, gender, Twitter use, and student status as covariates. Means not sharing a subscript significantly differ on pairwise comparisons ($p < .05$).
One common explanation for differences in credibility in the literature stems from differences in use. Thus, we explored whether individuals who reported using Twitter had different credibility perceptions between conditions, relative to non-users. To do so, we replicated the ANCOVA analysis reported earlier, but treated Twitter use as an additional moderating factor, rather than a covariate. The results gave no indication of a moderating role for Twitter use when considering either source credibility, $F(2, 182) = 0.17, \text{ ns}$; or message credibility, $F(2, 181) = 0.67, \text{ ns}$. Unfortunately, there are too few individuals in the sample who used Twitter to further probe for an influence based on differing levels or types of use.

Overall, then, the results suggest that Twitter is seen as a less credible source that may or may not present a less credible message, but that this might not affect opinions about the issue. The similar credibility for short and long stories suggests that level of information does not account for Twitter’s lower credibility. To further explore how the “location” in which a story appears relates to credibility, we carried out a second study holding the length and level of information more constant. In this study, we added a third potential “source,” a private blog page posting a story link. This allows us to more fully explore the outlets through which individuals might gain access to the same news: the outlet’s Web site, the official Twitter feed, and an unofficial blog reference. Given the unclear findings for information satisfaction and the removal of variance in level of information provided, we omit that variable in Study 2.

Study 2

Method

As in Study 1, participants were recruited by students at the same university, sent an e-mail directing them to the online study, and randomly assigned to one of three conditions: Twitter, newspaper Web site, or blog. Participants were more explicitly told that they were viewing a screen capture that might include only brief text, based on the belief that many of the participants who prematurely quit in Study 1 did so because they did not perceive the Twitter or short versions as a complete “story” and, thus, were reluctant to answer questions about the presented content. This appears to have been effective, as numbers across conditions were more equal. No participants had to be removed for incorrectly identifying the topic. The total $N$ was 435. Similar to the first study, participants were primarily current students (69.1%), women (62.1%), White (87.6%), and young (ages ranged from 18–68; $M = 26.31, SD = 11.05$). As with Study 1, a minority of participants (21.1%) used Twitter, but this rate exceeded estimates for overall population use. For this study, data were collected in the Spring of 2010.

Participants viewed the story description on The New York Times Web site, The New York Times’ Twitter feed on Twitter’s site, or on an individual’s blog created specifically for this study. In all conditions, the same story was used. This story appeared in the “Energy & Environment” section of The New York Times, with the
headline “Wind Power Grows 39% for the Year.” To minimize the potential confound of level of information, The New York Times version appeared on a page listing a number of different stories, giving only the headline and a one-sentence summary, with the byline of the author. The other stories listed were blurred. The Twitter version used the actual “nytimes” feed page, as in Study 1, with the original tweet visible (reading “Wind Power Grows 39% for the Year”), and followed by a bit.ly link and a note that it was posted recently. As with the version from the newspaper Web site, other stories were listed above and below this tweet, but they were blurred. Finally, the blog version appeared on a blog created just for this study using Google’s Blogger software. The blog was created under a fictitious username with the title “A day in the life . . .,” and some fictitious entries were created. The New York Times story was listed in a blog posting with the headline “NYTimes: Wind Power Grows 39% for the Year,” text saying “Check out this New York Times article on wind power,” the same image of windmills that appeared on the original New York Times Web site, and a link reading “New York Times article.” Other entries above and below this post were blurred. In total, 158 participants viewed The New York Times version, 131 viewed the Twitter version, and 146 viewed the blog version.

**Dependent Variables**

*Message credibility.* Participants were asked to evaluate the credibility of the message using the same two-item index as in Study 1 ($M = 4.22$, $SD = 1.39$; Cronbach’s $\alpha = .86$).

*Source credibility.* Participants were also asked to evaluate the credibility of the message source, again using the same two-item index as in Study 1 ($M = 4.36$, $SD = 1.55$; $\alpha = .93$).

*Issue importance.* Because analysis in the prior study showed little difference between personal and general importance, as well as immediate and future importance, in this study we relied on a single question to measure issue importance. Participants were asked how much they agreed with the statement, “This is an important issue,” using a 7-point agreement scale ($M = 4.96$, $SD = 1.41$).

*Information-seeking intent.* Participants were asked how likely they were to pursue this issue further through a variety of venues. An index was formed based on 11 questions, all of which were presented on a 7-point scale ranging from 1 (*not likely*) to 7 (*very likely*) ($M = 3.61$, $SD = 1.33$; $\alpha = .92$). Participants were asked about the same venues as Study 1, and because of the added blog condition, participants also indicated how likely they were to read comments about this issue and read blogs discussing this issue.

**Results**

Initial analyses were ANCOVA controlling for Twitter use, age, gender, and student status, with story condition as the independent variable. In Study 1, $H1a$ was not
supported, although the results were directionally consistent with expectations. Here, however, the results show a clear effect of story location on message credibility judgments, $F(2, 385) = 10.56, p < .001 (\eta^2_p = .05)$. As shown in Table 2, the story appearing directly on The New York Times Web page was seen as significantly more credible than stories linked from a blog or Twitter. The latter were not significantly different from one another, although the blog actually scored slightly higher on the scale, despite not having any official affiliation with The New York Times.

Data for H1b and judgments about source credibility are consistent with the previous result and the findings from Study 1. Story location had a significant effect on source credibility judgments, $F(2, 385) = 10.37, p < .001 (\eta^2_p = .05)$. As Table 2 shows, the findings are consistent with the results for message credibility, with scores for The New York Times significantly higher than the other outlets and the blog actually scoring slightly, albeit not significantly, higher than Twitter.

Despite the differences in credibility, Study 1 found no support for H2, involving issue importance and length. These analyses can only partly address this hypothesis, as only one story length was included. We can still evaluate whether the posting location for the story mattered, and these results do show an effect of story location on perceived importance, $F(2, 383) = 6.68, p < .01 (\eta^2_p = .03)$. As shown in Table 2, the results are consistent with the idea that Twitter posts would make the story seem less important. The New York Times story was associated with greater issue importance than the Twitter post, but the blog post actually led to the greatest perceived importance (albeit not significantly above the score for The New York Times).

Finally, the findings show no significant effect of story location on information-seeking intention, $F(2, 395) = 0.11, ns$. We carried out additional analyses to further explore this result. A regression model predicting information seeking with credibility and issue importance, along with the control variables, showed that message credibility ($\beta = .36, p < .05$) and issue importance ($\beta = .28, p < .001$) were both significant, positive predictors of information seeking. Once again, the results may be somewhat complicated by the strong correlation between the two measures of credibility ($r = .86, p < .001$), and zero-order correlation tests showed a significant, positive correlation of both message credibility ($r = .26, p < .001$) and source credibility

---

Table 2  Credibility, Importance, and Information-Seeking Scores by Post Location (Study 2)

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>Twitter</th>
<th>New York Times</th>
<th>Blog</th>
</tr>
</thead>
<tbody>
<tr>
<td>Message credibility</td>
<td>3.90&lt;sub&gt;a&lt;/sub&gt;</td>
<td>4.64&lt;sub&gt;b&lt;/sub&gt;</td>
<td>4.14&lt;sub&gt;a&lt;/sub&gt;</td>
</tr>
<tr>
<td>Source credibility</td>
<td>4.02&lt;sub&gt;a&lt;/sub&gt;</td>
<td>4.82&lt;sub&gt;b&lt;/sub&gt;</td>
<td>4.25&lt;sub&gt;a&lt;/sub&gt;</td>
</tr>
<tr>
<td>Importance</td>
<td>4.63&lt;sub&gt;a&lt;/sub&gt;</td>
<td>5.03&lt;sub&gt;b&lt;/sub&gt;</td>
<td>5.27&lt;sub&gt;b&lt;/sub&gt;</td>
</tr>
<tr>
<td>Information seeking</td>
<td>3.65&lt;sub&gt;a&lt;/sub&gt;</td>
<td>3.59&lt;sub&gt;a&lt;/sub&gt;</td>
<td>3.65&lt;sub&gt;a&lt;/sub&gt;</td>
</tr>
</tbody>
</table>

Note. Scores shown are estimated means with age, gender, Twitter use, and student status as covariates. Means not sharing a subscript significantly differ on pairwise comparisons ($p < .05$).
(r = .19, p < .001) with information seeking. Thus, similar to Study 1, it appears that story location may affect information-seeking intent indirectly—in this case, by influencing both credibility and importance judgments.

As in Study 1, we also carried out further analyses to consider why participants viewed Twitter as less credible, introducing Twitter use as a moderating factor instead of as a covariate. As in the first study, there was no significant interaction of Twitter use and story location in affecting either source credibility, $F(2, 383) = 0.40$, $ns$; or message credibility, $F(2, 383) = 0.41$, $ns$. Similarly, there was also no significant interaction involving these variables in predicting issue importance, $F(2, 381) = 0.45$, $ns$. In Study 1, we were unable to probe this further due to the small number of Twitter users. Thanks to the larger $N$ for this study, we were able to carry out exploratory analyses to more fully address the question of how Twitter use might affect credibility judgments. We considered the differences in credibility between the Twitter and newspaper conditions described earlier, focusing specifically on the 54 individuals in one of those conditions who indicated they used Twitter. Obviously, this gives us limited power, and we note results even approaching significance for that reason. We asked individuals who used Twitter to indicate (on a 7-point Likert scale) how much they agreed that “Twitter is a great way to get information” ($M = 4.03$, $SD = 1.90$), and “I follow a number of other users on Twitter” ($M = 4.90$, $SD = 1.92$).

We then carried out a series of regression analyses that included the control variables of age, gender, student status, a dummy-coded variable indicating whether individuals were in the Twitter or The New York Times conditions, one of the Twitter use/attitude measures, and an interaction term reflecting the product of that measure and the condition variable. The primary point of these analyses is that last term; a significant interaction would suggest that, although Twitter use itself did not moderate the findings, the specific amount or nature of use might. Overall, the results fall shy of significance, although the patterns are consistent with one another and with what one might expect if positive attitudes toward Twitter could enhance credibility. Indeed, if the analyses were run including the blog condition as well, many of these coefficients would be significant, but interpretation would be more suspect. Both interaction terms involving the measure “Twitter is a great way to get information” approach traditional significance levels. For source credibility, $b = .38$, $p = .08$; and for message credibility, $b = .33$, $p = .10$. In both cases, then, the result suggests for the participants in this study, at least, the gap between credibility in the Twitter and The New York Times conditions—which was negative, on average, reflecting lower credibility for Twitter—grew smaller as positive perceptions of Twitter increased. Yet, as the results show, the difference in credibility is so low at the lowest level of perceptions (the coefficient for the Twitter condition reflects the difference when perception scores equal 0) that even at the highest levels of perception, credibility between the two sources is essentially equal. It does not appear that even those most favorably disposed toward Twitter ever see it as notably more credible than the newspaper’s own Web site. Results showed a similar pattern for the variable measuring whether individuals follow others on Twitter, but were even further from
significance. For source credibility, the interaction of following and condition had $b = .25$, $p = .19$; and for message credibility had $b = .27$, $p = .13$.

In summary, it may be that in a study specifically of Twitter users, with greater power to detect relationships, active users of Twitter and those with especially favorable impressions of the medium would not display the same skeptical outlook found in the results as a whole. However, the data here are not strong enough to conclusively support that, and they show that when expressed in terms of mere use, those who are “on” Twitter do not show significantly more favorable responses to the medium than the majority of people online, who are non-users.

Discussion

Although research has considered differences in credibility among online information sources, it has generally not explored how the same information, from the same source, but distributed through different channels, would be perceived differently. In focusing on the use of Twitter by a news organization, we were able to consistently demonstrate lowered source credibility, although *The New York Times* is identified as the source of the information, and the Twitter post actually appears on the “official” *New York Times* feed page. We found a similar pattern for message credibility, although only in Study 2 was the difference significant. Some caution is necessary, as the net effect size in these results was fairly small and the sample was limited. The study offers good reason to question the credibility of media messages distributed by Twitter, but it would be premature to suggest that content shared this way suffers from a major credibility problem.

Nonetheless, this result provides a notable contrast to studies that show blogs have similar credibility to mainstream news. Whereas the “filtering” performed by a blog may suggest the information is worthwhile, Twitter apparently lacks this effect. Perhaps this stems from media coverage of Twitter linking it to celebrities and shallow posts from “average” Americans. Exploring exact perceptions of Twitter and their role in credibility judgments would be a worthy future direction. In this study, we did not want to prompt people to necessarily view the “source” as either Twitter or *The New York Times*, as the ambiguity of the source is a meaningful feature of media use of Twitter. By contrast, blogs more explicitly indicate both an originating and selecting source; a future study of a Twitter feed that serves a selecting function would be a useful comparison. Notably, however, in this study even the blog post was seen as less credible; and, although scores were higher than those of Twitter, the difference was not significant. Of course, whereas most widely read blogs would have a recognizable “brand” themselves, as well as ongoing discussion and further analysis, in this study we relied on a simple, direct link to the story—an expression of “I noticed this,” similar to what happens with Twitter. Thus, the blog created for this study probably scored worse in terms of credibility than a typical, widely known blog might, particularly among Internet-savvy individuals. Yet, the blog did not perform worse than Twitter, despite not having a recognizable brand or being presented as the
official feed of a major newspaper, suggesting that the results probably underestimate the degree to which Twitter posts may be viewed more negatively than blog links.

Meanwhile, these data do not fully illuminate the reasons for the differences. Prior studies have suggested that individuals who are regular consumers of a specific news source will see it as more credible. Yet, our analysis did not show a moderating effect of mere Twitter use. However, it could be that many of those saying they “used” Twitter did so only rarely and did not specifically follow news organizations. Further analyses in Study 2 were limited by the small number of Twitter users, but did offer hints that people with more favorable impressions or greater use of Twitter saw smaller credibility differences between sources, although not to the degree that they actually preferred Twitter. Further analysis in a study with greater numbers of Twitter users or with users of the specific feed being considered would be valuable.

This raises a broader question regarding the implications of these results in natural settings. On the one hand, it is tempting to dismiss the findings because a minority of participants made any use of Twitter. Yet, our study actually showed higher levels of use than those reported on representative surveys of Internet users. It is an open question whether “non-users” would actually be affected by Twitter content. For the most part, Twitter is a “subscription” driven service, with individuals receiving only those feeds they have deliberately chosen. However, some inadvertent exposure is possible. Google has at least intermittently provided Twitter results among search outcomes. Many blogs and other Web sites display a scrolling list of recent tweets that might be available even to those who do not see themselves as Twitter users. The New York Times recently moved to a subscription wall model that allows individuals to read a limited number of stories directly from the front page, but providing access to additional content when it is linked, as through tweets or blogs. This might prompt more people to initially encounter stories in those ways, even if they are not primarily users of those media. Moreover, we do not know whether greater use of Twitter is causing more positive credibility judgments. It could be that those with the greatest trust in Twitter gravitated toward it first, and later users will remain skeptical. Ultimately, the majority of Twitter users may differ from those represented in this study, but the relatively rare case of individuals first encountering a media outlet via Twitter, as reflected here, may be more meaningful. These individuals, after all, are the ones who will decide whether to subscribe to a feed or otherwise pursue further content. Nonetheless, future research should evaluate the sampling strategy carefully. The study used here relied on a student-heavy population, which actually helped boost the number of Twitter users, but compromises the ability to generalize to the population as a whole. A more representative sample would have fewer users but greater external validity, whereas a more deliberate sample of users would allow further probing of the effects of usage and perceptions of Twitter.

The study also explores the potential consequences of credibility judgments. We argued that Twitter should lead to stories seeming less important due to both low credibility and brevity. In Study 1, we did not find support for this; issue importance did not vary between conditions. In Study 2, however, the Twitter version was seen as less important. It may be that in Study 1, the issue was already too widely known and
discussed for a single manipulation to meaningfully affect judgments, making Study 2 a better test. If so, then the decreased credibility of Twitter-posted content has other meaningful effects. Not only are people more skeptical of the information, but they are also concluding that the underlying issues matter less. Notably, although the blog posting also suffered from credibility concerns, it was seen as more important than the Twitter post, suggesting that the added importance cue of having someone select this story from among many may have influenced perceived importance. This underscores the ambiguity of Twitter as a source. Although clearly perceived as distinct from content directly obtained from The New York Times, materials posted on The New York Times’ Twitter feed do not appear to benefit from the positive influence of coming from a “selecting” source.

Although credibility and importance are significant predictors of information-seeking intent, in both studies the manipulations failed to influence this third outcome variable. There may be some offsetting factor at play. Perhaps, as speculated, the tweet format created curiosity about the issue, but the effects on credibility and importance canceled this out; or, perhaps because participants were primarily college students, their willingness to pursue further news is too low, regardless of stimulus. That said, there are hints that the manipulation may have an indirect effect, as some aspects of credibility were correlated with information-seeking intent, and importance was a consistent, strong predictor of such intent. Perhaps repeated exposure to content on a media Web site would amplify information seeking relative to monitoring the same site via Twitter. Another consideration is the measurement of intended information-seeking behavior. Any such self-report data are speculative, and it is socially desirable to overestimate likely news use. Because of the software used to gather data, it was impossible to allow participants to actually click on the link in the various stories, but a design where this was possible would offer a useful measure.

Overall, however, the findings here offer important theoretical and practical insights. In contrast to many studies of online credibility, this study shows that even somewhat regular users of Twitter do not see it as providing more credible information, and the population as a whole is unusually skeptical of Twitter relative to other means of distribution. Twitter seems to elicit a negative reaction from many. The exact mechanisms are unclear, and we cannot say if this would carry over to other microblogging tools that may be under development. On the surface, however, it is noteworthy simply because unlike traditional blogs, Twitter here is not serving as a selecting source—the decision to share this story was made by the originating source, The New York Times. Yet, participants still viewed the content on Twitter differently, suggesting that the Twitter “brand” has some independent cue. Perhaps the positive responses to other selecting sources are also due to cues, and not to reasoned evaluations about the benefits of custom-selected material. At an applied level, this study suggests the need for caution in the use of Twitter as a way to distribute news. Despite the official New York Times “stamp,” these stories were still viewed in a more negative light when posted to Twitter. Participants even saw the larger news organization as less credible. It would be premature to say that using Twitter is necessarily
hurting *The New York Times*, but absent evidence showing a positive effect, skepticism seems warranted. The recent move by *The New York Times* to promote subscriptions appears mixed, from this perspective; however, because it came after this study was conducted, we cannot offer direct empirical insights on this point. The newspaper may be cultivating a more engaged, trusting set of core users, but it may also be increasing the chance that non-subscribers will come away with a negative impression of content if they encounter it through Twitter or some other means. At the very least, news organizations and others (such as corporations) should pay attention to continued research in this area and not reflexively embrace all new distribution technologies as equally beneficial.

**Note**

To save space, we present the full results for these analyses in this note, and we describe only a few critical findings. For source credibility, there are two models. The first includes the measure where participants indicated Twitter was a great way to get information, along with the relevant interaction term. Unstandardized coefficients and standard errors, in parentheses, in this model were as follows: age, $-0.06 (0.03)$; gender $-0.07 (0.38)$; student status, $0.60 (0.63)$; Twitter is a great source, $0.04 (0.15)$; Twitter condition, $-1.85 (1.04)$; great Source $\times$ Condition, $0.37 (0.21)$; and total $R^2 = .26$. The second model includes the measure where participants indicated following a number of other users on Twitter, along with the relevant interaction term. The coefficients and standard errors, in parentheses, for this model were as follows: age, $-0.06 (0.03)$; gender $-0.07 (0.38)$; student status, $0.60 (0.63)$; follow others, $0.04 (0.15)$; Twitter condition, $-1.85 (1.04)$; Follow $\times$ Condition, $0.37 (0.21)$; and total $R^2 = .26$. For message credibility, there were also two models. For the model with “Twitter is a great way to get information,” the coefficients and standard errors, in parentheses, were as follows: age, $-0.03 (0.02)$; gender $-0.05 (0.36)$; student status, $-0.15 (0.60)$; Twitter is a great source, $-0.04 (0.14)$; Twitter condition, $-1.84 (0.82)$; great Source $\times$ Condition, $0.33 (0.20)$; and total $R^2 = .16$. For the model with “I follow a number of other users,” the coefficients and standard errors, in parentheses, were as follows: age, $-0.05 (0.03)$; gender, $-0.29 (0.38)$; student status, $0.59 (0.61)$; follow others, $0.16 (0.13)$; Twitter condition, $-1.69 (0.99)$; Follow $\times$ Condition, $0.25 (0.18)$; and total $R^2 = .28$.

**References**


