

Aggregation, clickbait and their effect on perceptions of journalistic credibility and quality

Abstract

Many journalists and industry observers lament that aggregating news underneath sensational headlines will erode credibility and turn off readers. While some scholarly work has studied journalists' perspectives of this practice, little has been done to understand what audiences think of aggregation and clickbait. This study uses published original and aggregated news articles as stimuli in two online experiments to test readers' perceptions of news aggregation and clickbait. Aggregation itself has little effect on perceptions of credibility and quality; instead, writing proficiency is more closely linked to these perceptions. Results also suggest clickbait headlines may lower perceptions of credibility and quality.

Keywords: journalism, news aggregation, credibility, journalism quality, clickbait, online news

Authors:

Logan Molyneux
Assistant Professor, Journalism
Klein College of Media and Communication
Temple University
2020 N. 13th Street
Philadelphia, PA 19122
logan@temple.edu
<https://orcid.org/0000-0001-7382-3065>

Logan Molyneux is an assistant professor of journalism at Temple University. His research focuses on journalists' use of mobile and social media and changing journalistic norms.

Mark Coddington
Assistant Professor, Journalism
Washington & Lee University
coddingtonm@wlu.edu
<https://orcid.org/0000-0001-6664-6152>

Mark Coddington was a newspaper reporter in Nebraska before earning his Ph.D. in journalism from the University of Texas at Austin. His research focuses on media sociology, online news production, networked and participatory journalism, news aggregation, and social media and journalism.

(Both authors received their Ph.D. from The University of Texas at Austin.)

Funding for this study was provided by Temple University.

Aggregation, clickbait and their effect on perceptions of journalistic credibility and quality

Journalists have wrung their hands over many developments in their industry over the past decades, including economic instability, fragmentation of the audience, and incursions by new entrants in the field, to name a few. But one strain of these worries has wondered whether journalists themselves are making things worse: Are certain journalistic practices cheapening the product and driving audiences away? Two increasingly common practices have headlined this branch of the discussion, both designed to drive web traffic and attention: news aggregation, especially by reformulating and republishing news based on others' work, and the use of clickbait headlines.

News aggregation is defined by Coddington (2018) as the practice of taking information from published sources, reshaping it, and republishing it within a single place. The type of largely manual aggregation work done by journalists has become one of the dominant forms of digital newswork (Anderson, 2013a). This type of aggregation, which typically involves linking to and rewriting or commenting on work that has already appeared elsewhere, has commonly been associated with political blogs and social media-oriented sites like BuzzFeed (Carlson, 2007; Tandoc & Foo, 2018). But many of journalism's most traditionally respected news organizations have also added or expanded aggregation teams, with units such as *The New York Times*' Express Team (Sullivan, 2016) and *The Washington Post*'s breaking-news desk (Rieder, 2014) building on the work of others to produce news, or even aggregating their organization's own stories (Benton, 2018). (The form of automated search-based aggregation led by Google News has also figured prominently in journalists' and academics' understanding of the concept [Chyi, Lewis, & Zheng, 2016], though this study focuses specifically on the manually driven form of aggregating news within articles, rather than across them.) Journalists themselves have

often characterized aggregated news as one of the lowest rungs on their professional ladder, if it has qualified as journalism at all (Author, forthcoming). As aggregation has become a more established part of newswork, journalists' concerns over the quality and legitimacy of aggregation have persisted.

The growth of this form of aggregation has been fueled, at least in substantial part, by the goal of driving traffic by having the hottest stories, even if they were originally reported elsewhere. Even as criticism of aggregation crested in the late 2000s and early 2010s, the commercial pressures to pile up pageviews in order to feed a digital advertising-based business model pushed more news organizations to rely on aggregation themselves (Author, forthcoming). The same traffic-based economic model has also fed the rise of dubious methods of attracting attention to hastily produced online material. Most notably, 'clickbait' has emerged as a widely used term to deride and dismiss content that exists more as a way to lure audiences to click on it — inevitably letting those audiences down — than for any informative purpose (Hamblin, 2014; Klinger & McBride, 2016). Aggregation and clickbait are not inextricably connected; one is a form of information gathering and the other a means of drawing attention to information online. But they have both often functioned together as a strategy for drawing the type of large-scale traffic that an advertising-based economic model tends to require. Aggregation allows news organizations to cheaply and quickly produce content on popular topics, and packaging that aggregation as clickbait is aimed at maximizing the traffic such hastily produced, ephemeral content can draw (Author, forthcoming). Indeed, some journalists and media critics have closely associated this term with aggregated content (e.g., Barner, 2017; DiDomizio, 2018; Luckie, 2015; Owens, 2015), and used it as an indicator of both ethical and economic corner-cutting.

Undergirding much of this criticism, both of clickbait and of aggregation more generally, is journalists' argument that these practices, **which may produce short-term gains in web traffic, are ineffective in the long run as audiences tire of them. That is**, aggregating news underneath sensational headlines is a sure way to erode credibility and lose readers long-term as they view your content as uninformative, low-quality commodity news. This argument is often made explicitly, even by those who have practiced this type of journalism, as they insist that their work is of a higher form (Klinger & McBride, 2016; Luckie, 2015; Owens, 2015; Smith, 2014). But this argument has been made without consulting the audiences themselves — in fact, audiences' perceptions regarding aggregation and 'clickbait' content have received little empirical testing. This study aims to fill that gap by examining whether audiences observe the same distinctions in quality and credibility that journalists allege exist between aggregated and traditional newswriting, **and if perceptions of clickbait headlines lead to lower credibility for news articles, especially in combination with aggregation**. Aggregation has primarily been studied as a form of news production, with research illuminating its practitioners' attitudes toward it (Anderson, 2013a, 2013b; Boyer, 2013; Coddington, 2018; Vobič & Milojević, 2014). But scholars have done little to examine aggregation from the audience's perspective. This study brings in that perspective through an experiment testing readers' perceptions of credibility and quality regarding aggregated and originally reported news, as well as 'clickbait'-style headlines.

Literature review

News Aggregation

Aggregated news encompasses a wide variety of news forms, from email newsletters on niche subjects, to articles summarizing the buzz around the latest viral video, to mobile news apps that pull in feeds from numerous sources. **Following Coddington (2018), we define**

aggregation as journalists' work of taking information from published sources, reshaping it, and republishing it, often in an abbreviated form, within a single place. This definition could potentially include aggregators outside of the news media, but we are focused here on news aggregation — aggregation that deals with forming news accounts based at least in part on sources published elsewhere, and specifically the variety that aggregates information within the production of a single article, rather than aggregating many articles within a service, such as in Google News or Apple News.

As it's practiced in today's news organizations, aggregation often involves monitoring numerous streams of online information, then filtering, combining, repackaging, and recontextualizing it (Anderson, 2013a, 2013b; Boyer, 2013; Coddington, 2018). The end result often consists of short bites of information punctuated with links and embedded media such as videos and social media posts connecting the aggregated information with its published sources. Aggregation is hardly new in journalism — newspapers have been wantonly republishing each other's material since at least the 18th century (Kielbowicz, 1982) — and aggregation often blurs with reporting, both in the work involved and the content produced. But aggregation is emerging as a distinctive and increasingly widespread form of online news production. As Boyer (2013) observes, “newsmaking today is as much about managing multiple fast-moving flows of information already in circulation as it is about locating and sharing ‘new’ news” (p. 2).

Still, aggregation has often been viewed as a lesser form of journalism — less substantial, less valuable, and less reliable because of its greater distance from the evidence it uses (Coddington, 2018). Erik Neveu (2014) makes a particularly incisive version of this critique, arguing that today's compacted journalism inevitably oversimplifies complex events and strips out both depth and narrative quality. Neveu's distaste for aggregated news is echoed among

journalists. Journalistic rhetoric about aggregators throughout the last decade has been filled with language like “parasites” and “vampires” or “sleazy” (Arrington, 2010; Hagey, 2010). Because aggregation is derivative of other journalistic work, it treads close to journalism’s ethical boundaries as well. “Some folks have yet another word for aggregation and related activities on the Web,” wrote the prominent American editor Michael Kinsley. “They call it ‘plagiarism’” (2016). Aggregators have internalized these characterizations as well, referring to themselves as “copy-pasters,” “recyclers,” and “robots” who “pack news like sausages” (Vobič & Milojević, 2014, p. 1032).

To the extent that this discourse has taken place in public, one might expect it to have seeped into the public’s perception of aggregators as well. Initial indications support that idea. One recent poll found that only 38% of Americans had confidence that news aggregators were providing accurate and politically balanced news, lower than any other form of news it tested except for Internet-only news websites — many of which contain a fair amount of aggregation themselves (Gallup/Knight Foundation, 2018). And Choi and Kim (2017) found, paradoxically, that the more people used news aggregators, the less credible they found them — perhaps, the researchers concluded, because aggregators fed them repetitive news that wearied and frustrated them. This study aims to deepen these findings. Rather than measuring perception of aggregators’ reputation, we aim to measure perception of aggregated content itself, without regard to the professional baggage the term has accumulated.

Clickbait and Headlines

Headlines are not an inherent part of aggregation, but they have been taken on a particular connection to the culture and practice and aggregation through ‘clickbait’ — the term of derision that has emerged for attention-getting and often disappointing or somewhat deceiving online content. **Both aggregation and clickbait are used (often in tandem) as means of collecting**

clicks for material that may be of inferior quality. The tension of clickbait is rooted in headlines' two occasionally conflicting core purposes: To communicate vital information about the article to which they are attached and to grab the reader's attention (Andrew, 2007; Dor, 2003). Dor (2003) finds that journalists hold a tacit set of ideas about what makes an appropriate headline — short, clear, interesting, emphasizing new information and people with high news value, and connecting to prior events and assumptions. The result, researchers have found, are headlines that tend to exaggerate the tone of the attached article, especially negatively (Andrew, 2007; Ecker & Lewandowsky, 2014).

As headlines have shifted from print to an online context, many of those characteristics have remained. But two in particular have shifted: Clear and unambiguous headlines are no longer valued as highly as they once were, and the sense that headlines' first and most important payoff should be in reading the headline itself, rather than the accompanying article (Dor, 2003), has evaporated. In their place, economic incentives to draw as much traffic as possible have led journalists to focus on writing headlines whose primary goal is to maximize clicks, using techniques like A/B testing to ensure they do so (Hindman, 2015; Tandoc, 2014). One of the more effective tactics in attracting click-throughs have been headlines whose deliberate ambiguity tempt the reader to read the full article or whose brashness and undue certainty oversell the ambiguity of the actual news contained within, a practice commonly referred to as clickbait (Hamblin, 2014; Smith, 2014). In a study of online news about unverified claims, Silverman (2015) found that 13 percent of all articles contained this kind of misleading dissonance between headline and text, and a 2017 study found that use was high among mainstream media organizations, and especially in broadcast (Rony, Hassan, & Yousuf, 2017). In 2016, Facebook changed its News Feed algorithms to reduce the occurrence of these types of

‘clickbait’ headlines, which it defined as headlines that withhold information required to understand the story and create misleading expectations for the reader (Peysakhovich & Hendrix, 2016). **Clickbait headlines are often tied to news aggregation in practice — organizations such as BuzzFeed and Upworthy have relied heavily on pairing ambiguous or oversold headlines with their aggregated content — and in normative discussions. Recent research has found that professional norms chastising clickbait have emerged as a focal ethical point of consideration — if not a norm that is always followed — for many aggregators who work within the practice’s dominant model built around traffic and advertising (Author, forthcoming).**

Through its crest in popularity as a strategy in the 2010s, clickbait has been dogged by questions of its effectiveness. These headlines may entice clicks, the criticism goes, but they’re much worse at growing traffic through social sharing, and “they run the risk of fatiguing readers and eroding trust over time” (Klinger & McBride, 2016; see also Hindman, 2015; Smith, 2014). A 2016 study by Scacco and Muddiman provides some support for this idea, finding that question-based headlines produce more negative sentiments about the article to which they are attached, though “forward-reference” headlines such as “Why the American Economy is About to Boom Again,” which also withhold information to increase interest, had no effect. **Both Curry and Stroud (2016) and Kuiken et al. (2017), however, find that headlines with the characteristics of clickbait increase clicks, though the sentiment toward the article is not measured.** Clickbait headlines’ effectiveness in drawing clicks is rarely questioned; indeed, it is implied in the phenomenon’s name. But its effects on audience perception of the content it stands with are largely undetermined.¹

¹ Muddiman and Scacco (2019) have found that overheated headlines designed to stoke outrage decrease readers’ intended engagement with a news source and the credibility of that source after reading a headline and article, though those headlines do not withhold information as the ones in these studies and our own do.

Journalistic Credibility

While many consumers may not be fully aware of the reporting practices behind the news, research suggests that people form pointed opinions about the news they consume. Scholars have measured these opinions in terms of a message's credibility or quality (Thorson, Vraga, & Ekdale, 2010), the credibility of the journalist that produced them (Hovland & Weiss, 1951), or the credibility of the outlet or channel that carried them (Bucy, 2003; Johnson & Kaye, 2000). Cues about whether to trust a source or a medium are sometimes external to the message itself, as in the case of a recognizable brand, but are often embedded within it. Thus much credibility research commonly interprets credibility to mean believability and associates it strongly with idealized norms of journalistic practice, especially objectivity, as evidenced in journalistic products (Metzger et al., 2003). Credibility is a multidimensional concept and is frequently operationalized using measures of trustworthiness, accuracy, fairness, balance, absence of opinion, completeness, authenticity and believability (Gaziano & McGrath, 1986; Johnson & Kaye, 1998; Kiouisis, 2001; Kohring & Matthes, 2007; Appleman & Sundar, 2016). While research subjects may be asked to evaluate a source's accuracy, believability, etc., most have not ever met these sources, and their interactions with them are limited to the news messages they produce. Asking these questions about a medium draws not only from a particular message they consumed but from all relevant experience with that medium (explaining the important role of reliance in Johnson and Kaye's studies). Because of the important role that messages play in developing perceptions of credibility, credibility's importance is in its practical interest to journalists (Thorson et al., 2010). Audiences are assumed to develop trust (in sources and media) by consuming their messages, which trust would then lead them to have loyalty and consume more messages from that source or in that channel. Journalists therefore are cautioned

to craft messages that audiences will find credible, because credibility is the “currency of the realm” (Willis, 2009).

Attributes of the source, familiarity of the medium and features of the message itself have been shown to affect perceptions of credibility. Researchers have found that credibility is influenced by how much a person uses or relies on that particular medium (Johnson & Kaye, 2002, 2003). Structural attributes, including a balanced story presentation (Fico, Richardson, & Edwards, 2004) and sophisticated site design (Flanagin & Metzger, 2007) are also associated with perceived credibility. In addition, the technical quality of a message is positively associated with its perceived credibility (Cummins & Chambers, 2011). Together, these findings suggest that while messages may vary along several dimensions, it is both the qualities of the message itself and the user’s own experience with similar messages and sources that determine perceptions of credibility. This study manipulates message structure (aggregated vs. traditional reporting, clickbait vs. traditional headline) as the experimental condition and controls for characteristics of the user, including demographics, news consumption and news literacy.

Journalistic Quality

Journalism quality is, not surprisingly, often invoked by both journalists and audiences as a primary consideration in determining what kind of news to create and consume. But it remains a nebulous term in the public imagination and, to some degree, in communication research. For researchers, quality has overlapped heavily with credibility. In Sundar’s (1999) foundational study of audience perception of news, credibility and quality are two of the four major criteria for positive perception, along with liking, quality, and representativeness. Many of that study’s dimensions of credibility and quality are shared by studies measuring both message credibility (Chung, Kim, & Kim, 2010; Newhagen & Nass, 1989) as well as quality (Clerwall, 2014; Graefe, Haim, Haarmann, & Brosius, 2018).

Some of this ambiguity is inherent in the term. Quality is a particularly subjective characteristic; without a universally accepted normative base for journalistic quality, each individual's definition is heavily dependent on her distinct perspective. Urban and Schweiger (2014) identify six general dimensions that researchers have classified as normative criteria for news quality — diversity, relevance, accuracy, comprehensibility, impartiality, and ethics — and note that audiences' expectations for news quality largely reflect those criteria (e.g., Neuberger, 2014). But they also point out that when audiences are asked about what signals good journalism, many of their responses aren't nearly as sophisticated as the detail of those criteria might imply. And they find that when those criteria are varied in news articles, audiences make only small changes in their overall assessments of the article's quality.

This difficulty in recognizing and taking into account individual dimensions of news quality may be because people quickly scanning news articles use heuristic cues — such as media brands, uncivil comments beneath an article, or whether the article was written by a human — to assess quality in lieu of a full examination of the article's characteristics (Graefe et al., 2018; Prochazka, Weber, & Schweiger, 2018; Urban & Schweiger, 2014). Still, other substantive characteristics of news content, such as whether an article is edited or unedited (Vultee, 2016), **and whether an article was recommended by editors, other readers, or a computer (Sundar & Nass, 2001)**, have also been found to have significant effects on perceptions of quality. Clerwall (2014) and Graefe et al. (2016) both found that audiences saw few quality differences between automated and human-produced news articles, though they did tend to see the automated articles as more credible but drier. Within human production, the effect of the origin of the content produced — that is, originally reported versus aggregated secondhand — on perceived news quality has yet to be tested.

Existing research, then, has suggested that both aggregation and ambiguous clickbait headlines might be perceived as less credible and lower-quality by readers, but this suggestion requires explicit testing. Aggregation has long been perceived as lower-quality by journalists and, to a degree, by aggregators themselves, but it remains untested whether audiences share these views. Likewise with clickbait headlines, though some initial research has indicated that audiences might react more negatively to some kinds of clickbait headlines (Muddiman & Scacco, 2019; Scacco & Muddiman, 2016). This study thus poses the following hypotheses in order to empirically test audience perceptions of news aggregation and clickbait:

H1: Readers will rate aggregated news articles lower than original news articles in terms of (a) credibility and (b) quality.

H2: Readers will rate news articles with clickbait headlines lower than news articles with traditional headlines in terms of (a) credibility and (b) quality.

Methods

These hypotheses are tested in two studies, both using experiments. The first experiment employed a 2×2 (reporting style \times headline style) between-subjects factorial design, repeated with two versions of the stimuli. The second also used a 2×2 (reporting style \times writing proficiency) between-subjects factorial design, varying these elements in one set of stimuli. Participants ($N = 483$ and $N = 413$) were recruited via Amazon Mechanical Turk and invited to view the stimuli and fill out a questionnaire online. Methods and results from each study are presented separately.

Study 1 – Aggregation and clickbait as independent variables

Participants

Participants were recruited using Amazon's Mechanical Turk (MTurk) for the main experiment and a manipulation check of the stimuli. MTurk is an online platform that allows requesters to recruit workers and pay them for their time. The platform is commonly used to collect high-quality data for experiments (Hauser & Schwarz, 2016; Paolacci & Chandler, 2014; Peer, Vosgerau, & Acquisti, 2014), which do not require representative samples (Shadish, Cook, & Campbell, 2002). Participation was limited to adults living in the United States and workers who were approved for at least 95% of the tasks they completed. Workers were not allowed to re-take the study, and gender quotas were imposed to ensure that the sample was 49% male, 51% female. The average age for experiment participants was 39, ($N = 488$), and they spent an average of 10 minutes completing the study (reading a story and completing a questionnaire). The data was collected between Oct. 19 and 24, 2016.

Stimuli

Stimuli were prepared using real news articles published online. The selection criteria were that an original story had to have been aggregated by other news sites; that is, a story was published and subsequently aggregated and linked to by other sites. After identifying several such stories in August and September 2016, two specific examples were selected. These two examples both had an original story that included original reporting and were published on the website of a newspaper. Following this, they were both aggregated by several sources. The aggregated version used in this experiment represents aggregation done by a journalist at the organization where the aggregated story was published, rather than a newswire version of the story. In one case, the aggregation was considered to be high quality, and the other case was considered to be low-quality aggregation. In all cases, the story's publisher was hidden from participants, as publication source could be a potential confound. Only the body text was used.

Story A was about a boy, age 12, who enrolled at Cornell University as a freshman. An original version of the story was published in *The Washington Post* on August 15, 2016.² The reporter interviewed the parents, collected family photos, and spoke to sources at Cornell. The aggregated version of the story was published in *U.S. News & World Report* on August 19, 2016. This story does not use any original sources, but links to and quotes from the *Washington Post* story and a post on Texas Tech University's site (where the boy had enrolled in an independent study high school program)³. This aggregation was considered to be high quality because it pulled from multiple sources, followed traditional journalistic conventions (like the inverted pyramid and objectivity) and was generally well-written. The original version of the story was considerably longer than the aggregated version and thus had to be edited so that the content of the two stories was roughly equivalent.

Story B was about a Rhode Island man who used a potato gun to shoot corncobs at his neighbor's house following a dispute. The original version was published September 1, 2016, by *The Westerly Sun*, based in Pawcatuck, Connecticut. The reporter used court documents (apparently the arrest affidavit filed by police) and spoke with the police chief. The aggregated version of the story was published September 2, 2016, on Breitbart.com, a politically conservative news and opinion website. The story links to and quotes heavily from the *Westerly Sun* story with minimal efforts at rewriting or retelling the story. In fact, all sentences but one in the Breitbart story include partial quotes from the *Westerly Sun* story. This was considered to be low-quality aggregation. Neither version of the story had to be edited because they were nearly the same except for the abundant quotation marks and the link in the Breitbart story.

² See the appendix for links to these stories as they were originally published.

³ To preserve the look of the aggregated article and give readers contextual clues to its aggregated-ness, the stimuli included hyperlinks in all the same places as the aggregated articles. These links did not point anywhere, however, preventing participants from leaving the experimental setting by clicking a link in the stimulus.

New headlines were written for each story in a traditional style and to emulate clickbait. For Story A, the Cornell story, the traditional headline was “Cornell welcomes 12-year-old freshman”; the clickbait headline was “This new freshman at Cornell is how young?” For Story B, the potato gun story, the traditional headline was “Police: Westerly man shot corncobs at neighbor’s house,” the same as originally published by the *Sun*; the clickbait headline was “You won’t believe what this man shot at his neighbor’s house.”

Manipulation check

This study focuses on whether audiences notice and react to a distinction that professional journalists readily make: proximity to sources, defined in this experiment as either original reporting or aggregation. Because this difference may be subtle to many news consumers, a manipulation check was conducted involving 80 individuals from the target population (U.S. adults). The stories selected as stimuli are by definition “original” or “aggregated,” but the purpose of this manipulation check is to determine whether readers can notice this difference when prompted to look for it. Respondents were asked, for each pair of stories, which story contained more original reporting (“The reporter who wrote the story gathered the information him- or herself”) and which contained more aggregated news (“The reporter republished facts first published by other news organizations”). A significant majority of respondents said the original story had more original reporting (Story A: $\chi^2(1) = 12.5, p < .001$; Story B: $\chi^2(1) = 5.1, p < .05$). A majority of respondents said the aggregated stories had more aggregated news, but the difference was not statistically significant. These questions tested two sides of the same coin — which story was aggregated and which story was original. In each case, the audience recognized that the aggregated news story was not original, but did not necessarily perceive it as aggregated. This result may be expected given the blurriness of the boundary

between aggregation and reporting and the fact that the discourse surrounding this distinction occurs mostly in professional circles.

To test the effectiveness of the researcher-created headlines, respondents were asked which headline best fit the definition, “The headline was written primarily to entice you to click on it rather than to convey information.” Significantly more respondents chose the headline written to be clickbait (Story A: $\chi^2(1) = 6.2, p < .05$; Story B: $\chi^2(1) = 3.8, p = .05$). Finally, significant majorities of respondents said the stories in each pair were about the same length and contained about the same information.

Procedure

Following recruitment on MTurk, participants were taken to a questionnaire created using the online survey software Qualtrics. After viewing informed consent documentation, participants were instructed to read a news story. Participants were randomly assigned to one of the four conditions: (1) the original story with a traditional headline, (2) the original story with a clickbait headline, (3) the aggregated story with a traditional headline, or (4) the aggregated story with a clickbait headline. (These are numbered roughly in order of expected journalistic quality, valuing original sources over secondary ones and informational headlines over gimmicks.) After viewing the assigned version of Story A, participants were asked to rate it on two dimensions: credibility and quality. The process was repeated for Story B. Finally, respondents answered demographic questions and questions about their news consumption and news literacy.

Measures

The independent variables in this study were reporting style (original sources vs. aggregation) and headline style (traditional vs. clickbait). The study then measured credibility and quality as dependent variables. These dependent variables were measured twice, once for

each story. Demographics, online news consumption and news literacy were measured as control variables.

Credibility. Participants were asked to rate their agreement with the following statements about the article on a five-point scale: “can be trusted,” “is fair,” “is informative,” “is accurate,” “tells the whole story,” “is biased,” and “is opinionated.” The last two items were reverse coded. These items were adapted from the Gaziano & McGrath (1986) scale, as also used in other works (see Bucy, 2003; Johnson & Kaye, 2000). Responses were summed and divided by the number of items in the scale to produce an index of story credibility ranging from 1 to 5 (Story A: $M = 3.96$, $SD = .63$, $\alpha = .83$; Story B: $M = 3.59$, $SD = .67$, $\alpha = .83$; all credibility items together: $\alpha = .87$).

Quality. Participants were asked to rate their agreement with the following statements about the article on a five-point scale: “considers different opinions,” “reports relevant facts,” “is complete,” “is objective,” “does not contain errors or contradictions,” “is understandable,” “does not insult or discriminate.” These items were adapted from Prochazka and colleagues (2018). Responses were summed and divided by the number of items in the scale to produce an index of story quality ranging from 1 to 5 (Story A: $M = 3.93$, $SD = .49$, $\alpha = .67$; Story B: $M = 3.62$, $SD = .60$, $\alpha = .73$; all quality items together: $\alpha = .77$).

Online news consumption. Because of this study’s presentation online and its focus on differences in online news presentation, online news consumption was used as a control variable. Respondents were asked how often, on a five-point scale, they get news from the following sources: Online news sites, and news aggregators (Google News, Flipboard, Apple News). Respondents were also asked how often they get news on the following platforms: smartphone app or browser, tablet app or browser, and computer web browser. Responses were summed and

divided by the number of items to produce a scale of online news consumption ranging from 1 to 5 ($M = 2.47$, $SD = .82$, $\alpha = .70$).

News literacy. Though news literacy might be operationalized in many different ways, this study developed new items that measured a participant's facility with and interest in identifying sources behind the news. This is because the primary difference between aggregated news and original reporting, the main independent variable in this study, is a difference in sourcing. Accordingly, participants were asked to rate their agreement with the following statements on a five-point scale: "I pay attention to who produced a news story before I read, watch or listen," "I won't read a news story if it comes from a source I don't recognize," "News stories are usually the same no matter where they come from," "Who reports the news is as important to me as what the news is." The third item was reverse coded. Responses were summed and divided by the number of items in the scale to produce an index of story credibility ranging from 1 to 5 ($M = 3.66$, $SD = .94$, $\alpha = .72$).

Results

H1 proposed that readers will rate aggregated news articles lower than original news articles in terms of (a) credibility and (b) quality. One-way ANOVA tests controlling for headline condition found significant differences in credibility and quality ratings as a main effect of whether participants saw the original or aggregated versions of stories (see Table 1). For Story A, the aggregated version of the story had significantly higher credibility ($F(1, 483) = 5.79$, $p < .05$) and significantly higher quality ($F(1, 483) = 4.94$, $p < .05$) than the original. For Story B, the aggregated version of the story had significantly lower credibility ($F(1, 483) = 8.60$, $p < .01$) and quality ($F(1, 483) = 12.14$, $p < .001$) than the original.

H2 proposed that readers will rate news articles with clickbait headlines lower than news articles with traditional headlines in terms of (a) credibility and (b) quality. One-way ANOVA tests controlling for reporting style found significant differences as a main effect of whether participants saw a clickbait headline or a traditional one. For Story A, the version that appeared with a clickbait headline had significantly less quality ($F(1, 483) = 6.405, p < .05$), but there was no significant difference in credibility based on the clickbait headline (see Table 1). For Story B, the version that appeared with a clickbait headline was rated lower in terms of quality, but this difference only approached significance ($F(1, 483) = 2.83, p < .10$).

Interaction effects show a mixed picture. Journalistic theory would predict that the best case scenario (original reporting with a traditional headline) would have greater perceived quality and credibility than the worst case scenario (aggregated story with a clickbait headline) but the results do not show a clear pattern. That is, while perceptions of credibility and quality varied among the four conditions, they did not do so consistently, suggesting that these variations are based on something other than aggregation and clickbait.

Study 2 – Aggregation and writing proficiency as independent variables

The results of Study 1 suggest that aggregation itself was not a key differentiator for story credibility and quality. Instead, it appeared that evaluations of credibility and quality were more dependent on the writing proficiency of the story itself. In order to explicitly control for this possibility, a second experiment was designed to vary writing proficiency as well as reporting style (aggregated or original reporting). Choosing real news stories as stimuli benefits external validity, but the goal of this second study was to control the content of the stimuli to benefit internal validity. Study 2 followed the same procedures and recruited participants in the same way as Study 1, and it also employed the same measures. The only difference in the two

experiments was the stimuli, which in Study 2 were created by the researchers using an article originally published on Forbes.com about a breakthrough in semiconductor technology. Again, the publication source was not shown to participants. Data for Study 2 was collected between May 15 and 17, 2018.

Stimuli

The Forbes.com article used to create stimuli for this experiment was selected because of its relative obscurity in terms of subject matter, minimizing the chance that participants' existing knowledge of the story could act as a confound. Additionally, the story was written in a traditional journalistic style using original sources. This original version of the story, with minimal edits for length, was used as the original, clean version of the stimulus. Using this as a baseline, the researchers varied reporting style by changing all original sources to refer to other, previously published sources ("according to Forbes," "Forbes reported that," "IBM wrote in a blog that," and so on). Writing proficiency was varied by introducing errors into the copy (such as awkward syntax, long quotes, changing the order of the story, and using technical terms without any explanation). This was meant to simulate the sometimes haphazard nature of news aggregation on the internet. A manipulation check showed that the manipulated versions of the story included significantly less original reporting, significantly more aggregation, and were not as well-written as the original, while including about the same information. Participants were randomly assigned to view one of the four versions of the stimuli and answered the same battery of questions about credibility and quality, along with the same controls as in Study 1.

Results

Results of one-way ANOVA tests found significant differences only for the poorly written version of the story, and none for the aggregated version of the story (See Table 2). The aggregated version of the story was judged to be about as credible and have about the same

quality as the original. The poorly written version, on the other hand, was rated significantly lower in terms of credibility ($F(1, 413) = 14.25, p < .001$) and quality ($F(1, 413) = 26.50, p < .001$).

Discussion

As journalists have adopted practices of online news aggregation and clickbait as traffic-boosting strategies, one of the concerns within the profession has been that these forms of news are of lower quality than traditional reporting and therefore damage journalism's credibility. While this normative claim has been leveled at both aggregation and clickbait, often in tandem, this study found that audience perceptions of each practice vary.

Results from this experiment suggest that audiences aren't put off by the aggregation itself, and indeed, may not even notice it. In Study 1, aggregation produced inconsistent effects on quality and credibility, and in Study 2, which controlled content even more closely, aggregation's effects were non-significant. These results suggest that aggregation may not be perceived as negatively as many journalists fear — when done well, it may even be regarded more highly than original reporting. The key for journalists, then, as aggregation practices develop and become more widespread, will be to identify and adopt best practices such as adding contextual information and valuing creative writing and presentation, enabling them to publish work the public appreciates. Many of these norms are long established in the case of traditional news reporting, though the speed at which aggregation is produced and its professional marginality may have slowed their adoption among aggregators. Nonetheless, it appears that some of these norms carry substantial importance in audiences' perceptions of aggregated news reports as well as originally reported ones.

The divergence in significant results in Study 1 is presumably because of differences in writing proficiency among the articles used as stimuli. The aggregated version of Story A, about the young student at Cornell, is engaging and follows journalistic conventions of brevity, simplicity and a conversational tone. In short, it “reads well,” as editors are fond of saying.⁴ The aggregated version of Story B, on the other hand, links to one original report and quotes liberally from it in almost every sentence, essentially cutting up and stitching together the original without much care for readability. The story represents the type of shoddy, quick-hit work that concerns professional journalists and industry observers; it relies heavily on a single source of others’ work while adding little context or value of its own (Buttry, 2012; Martin, 2014). Thus it appears that other elements of production beyond the fact of aggregation (for instance, how well the story was written), account for the observed differences in measures of quality and credibility. Consequently, Study 2 was designed to control for writing proficiency, and in this case, aggregation had no measurable effect on credibility or quality. What instead led to significantly lower ratings of quality and credibility was lowering the writing proficiency of the story.

This study’s second hypothesis focused on clickbait, defined here as the practice of writing headlines that withhold information or oversell the story in order to entice readers to click through. Again, while metajournalistic discourse derides clickbait headlines as damaging to the news product they announce, readers’ attitudes may be more muted. Stories with clickbait headlines in this study were consistently judged to be of lower quality and lower credibility, but these differences were significant in only one case (the clickbait version of Story A had significantly lower credibility than the original; see Table 1). This suggests that the effect of

⁴ To put it more quantitatively, the original story has longer sentences (19.4 words per sentence compared with 16.2 in the aggregated version), longer paragraphs (37.1 words per paragraph compared with 26.9) and uses slightly longer words, on average (4.69 letters per word compared with 4.65). The original version is longer overall, containing 408 words compared with the aggregated version’s 323.

clickbait headlines on story credibility and quality is negative but rather small — small enough that this study did not have enough statistical power to detect it at a level of significance. Scacco and Muddiman’s (2016) findings were also marginally negative, but a follow-up study (Muddiman & Scacco, 2019) has found more negative and less ambiguous reactions to overblown headlines. While the differences observed here may be small in isolation, it’s possible that negative perceptions could develop over time. For instance, an organization that develops a reputation of writing clickbait headlines may come to be judged as less credible than one that employs the practice sparingly.⁵ The results presented here suggest caution in the use of clickbait headlines.

This study made use of actual news articles, a research design choice that was made with awareness of the benefits and drawbacks inherent in it. Because the stimuli were published news articles, Study 1 represents a case of high external validity, something experiments often lack. On the other hand, the nature of aggregated news required that the stories be circulated widely, increasing the chance that study participants had previously encountered the stories on their own. This possibility likely had little effect on the results, however, as the participants were randomly assigned and asked to rate the particular story they were shown. The questionnaire did not ask about content recall or familiarity. Finally, experiments such as these are not intended to produce generalizable results but to establish connections between a stimulus and a response. Thus the findings may not apply in every journalistic context.

Overall, these findings temper the anxiety journalists and media observers have expressed over news aggregation as an element of professional journalism practice. While audiences may consider the origin of news sources as a factor in determining a news story’s credibility or

⁵ This was the case with Upworthy, which became so synonymous with clickbait headlines that its co-founder apologized for “unleash[ing] a monster” as the company pivoted toward originally produced videos (Bilton, 2016).

quality, they appear to give significant weight to other elements of story production as well.

Clickbait headlines, on the other hand, appear to have broadly negative impact on audience perceptions, though the effects are small. Journalists are right to be concerned, then, about giving audiences the best quality product, but it appears that even aggregated stories, if done well, can be well-received.

Appendix A: Stimuli Source Material

Note: Text from the links below was used to create stimuli for these studies. Only the text was used and was presented without any indications of where it was originally published.

Study 1

Story A – Original

Available at: <https://www.washingtonpost.com/news/grade-point/wp/2016/08/15/a-12-year-old-is-off-to-the-ivy-league-its-a-challenge-to-keep-him-challenged-his-dad-says/>

Story A – Aggregated

Available at: <http://www.usnews.com/education/best-colleges/articles/2016-08-19/meet-jeremy-shuler-12-year-old-cornell-freshman>

Story B – Original

Available at: https://www.thewesterlysun.com/news/police-westerly-man-shot-corn-cobs-at-neighbor-s-house/article_deef0724-791c-5731-b124-22fc74fc4b6c.html

Story B – Aggregated

Available at: <http://www.breitbart.com/big-government/2016/09/02/man-arrested-shooting-corn-cobs-neighbors-house-homemade-gun/>

Study 2

Original story

Available at: <https://www.forbes.com/sites/aarontilley/2017/06/05/ibm-5nm-chips/>

References

Author, Forthcoming

Anderson, C. W. (2013a). *Rebuilding the news: Metropolitan journalism in the digital age*.

Philadelphia: Temple University Press.

Anderson, C. W. (2013b). What aggregators do: Towards a networked concept of journalistic expertise in the digital age. *Journalism*, 14, 1008-1023.

<https://doi.org/10.1177/1464884913492460>

Andrew, B. C. (2007). Media-generated shortcuts: Do newspaper headlines present another roadblock for low-information rationality? *The Harvard International Journal of Press/Politics*, 12(2), 24-43. <https://doi.org/10.1177/1081180X07299795>

Appelman, A., & Sundar, S. S. (2016). Measuring Message Credibility. *Journalism & Mass Communication Quarterly*, 93(1), 59–79. <https://doi.org/10.1177/1077699015606057>

Arrington, M. (2010, February 2). Everybody forgets the readers when they bash news aggregators. *TechCrunch*. Retrieved from <http://techcrunch.com/2010/02/02/everybody-forgets-the-readers-when-they-bash-news-aggregators/>

Barner, K. (2017, April 13). The viral-media editor who says cute babies and animals are her ‘bread and butter.’ *New York*. Retrieved from <http://nymag.com/thejob/2017/04/the-viral-media-editor-who-will-get-a-click-with-cute-babies-and-animals.html>

Benton, J. (2018, Oct. 2). Why The New York Times TL;DR’d its own 14,218-word Trump investigation. *Nieman Journalism Lab*. Retrieved from <http://www.niemanlab.org/2018/10/why-the-new-york-times-tldr-its-own-13000-word-trump-investigation/>

Bilton, R. (2016, April 13). A year into its new original content strategy, Upworthy is focusing on do-good videos instead of clickbait. *Nieman Journalism Lab*. Retrieved from

<http://www.niemanlab.org/2016/04/a-year-into-its-new-original-content-strategy-upworthy-is-focusing-on-do-good-videos-instead-of-clickbait/>

Boyer, D. (2013). *The life informatic: Newsmaking in the digital era*. Ithaca, NY: Cornell University Press.

Bucy, E. P. (2003). Media credibility reconsidered: Synergy effects between on-air and online news. *Journalism & Mass Communication Quarterly*, 80, 247-264.

<https://doi.org/10.1177/107769900308000202>

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science*, 6, 3-5.

<https://doi.org/10.1177/1745691610393980>

Buttry, S. (2012, May 16). Aggregation guidelines: Link, attribute, add value. *The Buttry Diary* [Blog post]. Retrieved from <https://stevebuttry.wordpress.com/2012/05/16/aggregation-guidelines-link-attribute-add-value/>

Carlson, M. (2007). Blogs and journalistic authority: The role of blogs in US Election Day 2004 coverage. *Journalism Studies*, 8, 264-279. <https://doi.org/10.1080/14616700601148861>

Choi, S., & Kim, J. (2017). Online news flow: Temporal/spatial exploitation and credibility. *Journalism*, 18, 1184-1205. <https://doi.org/10.1177/1464884916648096>

Chung, C. J., Kim, H., & Kim, J. H. (2010). An anatomy of the credibility of online newspapers. *Online Information Review*, 34, 669-685. <https://doi.org/10.1108/14684521011084564>

Chyi, H. I., Lewis, S. C., & Zheng, N. (2016). Parasite or partner: Coverage of Google News in an era of news aggregation. *Journalism & Mass Communication Quarterly*, 93, 789-815.

<https://doi.org/10.1177/1077699016629370>

Clerwall, C. (2014). Enter the robot journalist: Users' perceptions of automated content.

- Journalism Practice*, 8, 519-531. <https://doi.org/10.1080/17512786.2014.883116>
- Coddington, M. (2018). Gathering evidence of evidence: News aggregation as an epistemological practice. *Journalism* [Online before print]. <https://doi.org/10.1177/1464884918817608>
- Cummins, R. G., & Chambers, T. (2011). How production value impacts perceived technical quality, credibility, and economic value of video news. *Journalism & Mass Communication Quarterly*, 88, 737-752. <https://doi.org/10.1177/107769901108800404>
- Curry, A., & Stroud, N. J. (2016). Writing solutions headlines. *Engaging News Project*. Retrieved from <https://engagingnewsproject.org/research/solutions-headlines/>
- DiDomizio, N. (2018, Dec. 5). Just some reflections on my time at Mic. *The Nicolas Blog*. Retrieved from <https://thenicolasblog.com/2018/12/05/just-some-reflections-on-my-time-at-mic/>
- Dor, D. (2003). On newspaper headlines as relevance optimizers. *Journal of Pragmatics*, 35, 695-721. [https://doi.org/10.1016/S0378-2166\(02\)00134-0](https://doi.org/10.1016/S0378-2166(02)00134-0)
- Ecker, U. K. H., & Lewandowsky, S. (2014). The effects of subtle misinformation in news headlines. *Journal of Experimental Psychology: Applied*, 20, 323-335. <https://doi.org/10.1037/xap0000028>
- 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, 301-318. https://doi.org/10.1207/s15327825mcs0703_3
- 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, 319-342. <https://doi.org/10.1177/1461444807075015>

- Gallup/Knight Foundation (2018). American views: Trust, media and democracy. Retrieved from <https://knightfoundation.org/reports/american-views-trust-media-and-democracy>
- Gaziano, C., & McGrath, K. (1986). Measuring the concept of credibility. *Journalism and Mass Communication Quarterly*, 63, 451-462.
- Graefe, A., Haim, M., Haarmann, B., & Brosius, H.-B. (2018). Perception of automated computer-generated news: Credibility, expertise, and readability. *Journalism*, 19, 595-610. <https://doi.org/10.1177/1464884916641269>
- Hagey, K. (2010, September 23). Huffington vs. Downie on ‘parasites.’ *Politico*. Retrieved from <http://www.politico.com/news/stories/0910/42641.html>
- Hamblin, J. (2014, November 11). It’s everywhere, the clickbait. *The Atlantic*. Retrieved from <http://www.theatlantic.com/entertainment/archive/2014/11/clickbait-what-is/382545/>
- Hauser, D. J., & Schwarz, N. (2015). Attentive Turkers: MTurk participants perform better on online attention checks than do subject pool participants. *Behavior Research Methods*, 48, 400-407. <https://doi.org/10.3758/s13428-015-0578-z>
- Hindman, M. (2015). Stickier news: What newspapers don't know about web traffic has hurt them badly – but there is a better way. Shorenstein Center on Media, Politics and Public Policy, Discussion Paper D-93. Retrieved from <http://shorensteincenter.org/wp-content/uploads/2015/04/Stickier-News-Matthew-Hindman.pdf>
- Hovland, C. I., & Weiss, W. (1951). The influence of source credibility on communication effectiveness. *Public Opinion Quarterly*, 15, 635-650. <https://doi.org/10.1086/266350>
- Johnson, T. J., & Kaye, B. K. (1998). Cruising is believing?: Comparing Internet and traditional sources on media credibility measures. *Journalism & Mass Communication Quarterly*, 75, 325-340. <https://doi.org/10.1177/107769909807500208>

- Johnson, T. J., & Kaye, B. K. (2000). Using is believing: The influence of reliance on the credibility of online political information among politically interested Internet users. *Journalism & Mass Communication Quarterly*, 77, 865-879.
<https://doi.org/10.1177/107769900007700409>
- Johnson, T. J., & Kaye, B. K. (2002). Webelievability: A path model examining how convenience and reliance predict online credibility. *Journalism & Mass Communication Quarterly*, 79, 619-642. <https://doi.org/10.1177/107769900207900306>
- Johnson, T. J., & Kaye, B. K. (2003). A boost or bust for democracy?: How the Web influenced political attitudes and behaviors in the 1996 and 2000 presidential elections. *The Harvard International Journal of Press/Politics*, 8(3), 9-34.
<https://doi.org/10.1177/1081180X03008003002>
- Kielbowicz, R. B. (1982). Newsgathering by printers' exchanges before the telegraph. *Journalism History*, 9(2), 42-48.
- Kinsley, M. (2016, September). Donald Trump and the limits of data journalism. *Vanity Fair*. Retrieved from <http://www.vanityfair.com/news/2016/08/donald-trump-and-the-limits-of-data-journalism>
- Kiousis, S. (2001). Public trust or mistrust? Perceptions of media credibility in the information age. *Mass Communication and Society*, 4, 381-403.
https://doi.org/10.1207/S15327825MCS0404_4
- Klinger, L., & McBride, K. (2016, February 22). Stop calling every news article clickbait. *Poynter*. Retrieved from <http://www.poynter.org/2016/clickbait/397841/>
- Kohring, M., & Matthes, J. (2007). Trust in news media: Development and validation of a multidimensional scale. *Communication Research*, 34, 231-252.

<https://doi.org/10.1177/0093650206298071>

Kuiken, J., Schuth, A., Spitters, M., & Marx, M. (2017). Effective headlines of newspaper articles in a digital environment. *Digital Journalism*, 5, 1300-1314.

<https://doi.org/10.1080/21670811.2017.1279978>

Luckie, M. S. (2015, December). Adele and the death of clickbait. *Nieman Journalism Lab*.

Retrieved from <http://www.niemanlab.org/2015/12/adele-and-the-death-of-clickbait/>

Martin, F. (2014). The case for curatorial journalism ... or, can you really be an ethical aggregator? In L. Zion & D. Craig (Eds.), *Ethics for digital journalists: Emerging best practices* (pp. 87-102). New York: Routledge.

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. <https://doi.org/10.1080/23808985.2003.11679029>

Muddiman, A., & Scacco, J. (2019). Clickbait content may not be click-worthy. *Center for Media Engagement*. Retrieved from <https://mediaengagement.org/research/clickbait-content-may-not-be-click-worthy/>

Neuberger, C. (2014). The journalistic quality of Internet formats and services: Results of a user survey. *Digital Journalism*, 2, 419-433. <https://doi.org/10.1080/21670811.2014.892742>

Neveu, E. (2014). Revisiting narrative journalism as one of the futures of journalism. *Journalism Studies*, 15, 533-542. <https://doi.org/10.1080/1461670X.2014.885683>

Newhagen, J., & Nass, C. (1989). Differential criteria for evaluating credibility of newspapers and TV news. *Journalism Quarterly*, 66, 277-284.

Owens, S. (2015, January 5). 2015 is the year viral content becomes even more worthless. [Blog

post]. Retrieved from <http://www.simonowens.net/2015-is-the-year-viral-content-becomes-even-more-worthless>

Paolacci, G., & Chandler, J. (2014). Inside the Turk: Understanding Mechanical Turk as a participant pool. *Current Directions in Psychological Science*, 23, 184-188.

<https://doi.org/10.1177/0963721414531598>

Peysakhovich, A., & Hendrix, K. (2016, Aug. 4). Further reducing clickbait in Feed. *Facebook*.

Retrieved from <https://newsroom.fb.com/news/2016/08/news-feed-fyi-further-reducing-clickbait-in-feed/>

Prochazka, F., Weber, P., & Schweiger, W. (2018). Effects of civility and reasoning in user comments on perceived journalistic quality. *Journalism Studies*, 19, 62-78.

<https://doi.org/10.1080/1461670X.2016.1161497>

Rieder, R. (2014, April 23). The resurgent 'Washington Post.' *USA Today*. Retrieved from

<http://www.usatoday.com/story/money/columnist/rieder/2014/04/23/things-looking-up-at-washington-post/8046347/>

Rony, M. M. U., Hassan, N., & Yousuf, M. (2017). Diving deep into clickbaits: Who use them to what extents in which topics with what effects? *Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2017 (ASONAM '17)*, 232-239.

Scacco, J., & Muddiman, A. (2016). Investigating the influence of "clickbait" headlines.

Engaging News Project. Retrieved from <http://engagingnewsproject.org/research/clickbait-headlines/>

Silverman, C. (2015). *Lies, damn lies, and viral content: How news websites spread (and*

debunk) online rumors, unverified claims, and misinformation. New York: Tow Center for

- Digital Journalism. Retrieved from http://towcenter.org/wp-content/uploads/2015/02/LiesDamnLies_Silverman_TowCenter.pdf
- Smith, B. (2014, November 6). Why BuzzFeed doesn't do clickbait. *BuzzFeed*. Retrieved from <http://www.buzzfeed.com/bensmith/why-buzzfeed-doesnt-do-clickbait#.qhwQ3gKOzA>
- Sullivan, M. (2016, February 6). Sweetheart, get me readers. *The New York Times*. Retrieved from http://www.nytimes.com/2016/02/07/public-editor/new-york-times-express-team-margaret-sullivan-public-editor.html?_r=0
- Sundar, S. S. (1999). Exploring receivers' criteria for perception of print and online news. *Journalism & Mass Communication Quarterly*, 76, 373-386.
<https://doi.org/10.1177/107769909907600213>
- Sundar, S. S., & Nass, C. (2001). Conceptualizing sources in online news. *Journal of Communication*, 51, 52-72. <https://doi.org/10.1111/j.1460-2466.2001.tb02872.x>
- Tandoc Jr., E. C. (2014). Journalism is twerking? How web analytics is changing the process of gatekeeping. *New Media & Society*, 16, 559-575.
<https://doi.org/10.1177/1461444814530541>
- Tandoc Jr., E. C., & Foo, C. Y. W. (2018). Here's what BuzzFeed journalists think of their journalism. *Digital Journalism*, 6, 41-57. <https://doi.org/10.1080/21670811.2017.1332956>
- Thorson, K., Vraga, E., & Ekdale, B. (2010). Credibility in context: How uncivil online commentary affects news credibility. *Mass Communication and Society*, 13, 289-313.
<https://doi.org/10.1080/15205430903225571>
- Urban, J., & Schweiger, W. (2014). News quality from the recipients' perspective. *Journalism Studies*, 15, 821-840. <https://doi.org/10.1080/1461670X.2013.856670>
- Van Deursen, A. J. A. M., Helsper, E. J., & Eynon, R. (2016). Development and validation of the

Internet Skills Scale (ISS). *Information, Communication & Society*, 19, 804-823.

<https://doi.org/10.1080/1369118X.2015.1078834>

Vobič, I., & Milojević, A. (2014). “What we do is not actually journalism”: Role negotiations in online departments of two newspapers in Slovenia and Serbia. *Journalism*, 15, 1023-1040.

<https://doi.org/10.1177/1464884913511572>

Vultee, F. (2015). Audience perceptions of editing quality: Assessing traditional news routines in the digital age. *Digital Journalism*, 3, 832-849.

<https://doi.org/10.1080/21670811.2014.995938>

Willis, J. (2009). *The mind of a journalist: How reporters view themselves, their world, and their craft*. Thousand Oaks, CA: Sage.

Table 1. One-Way ANOVAs comparing ratings of credibility and quality in Study 1.

		<u>Credibility</u>	<u>Quality</u>
	<i>N</i>	Difference in mean score (vs. original)	Difference in mean score (vs. original)
Story A			
Aggregated	236	.138*	.098*
Clickbait	236	-.063	-.111*
Story B			
Aggregated	241	-.180**	-.187***
Clickbait	246	-.050	-.090 [†]

*Note: Total N = 483. Models included age, income, education, online news consumption, and news literacy as control variables. [†] p < .10, * p < .05, ** p < .01, *** p < .001.*

Table 2. One-Way ANOVAs comparing ratings of credibility and quality in Study 2.

		<u>Credibility</u>	<u>Quality</u>
	<i>N</i>	Difference in mean score (vs. original)	Difference in mean score (vs. original)
Story C			
Aggregated	211	-.019	.035
Messy	210	-.246***	-.268***

*Note: Total N = 413. Models included age, income, education, online news consumption, and news literacy as control variables. * $p < .05$, ** $p < .01$, *** $p < .001$.*