Is Automated Journalistic Writing Less Biased? An Experimental Test of Auto-Written and Human-Written News Stories

Yanfang Wu

Routledge
Taylor & Francis Group

Check for updates

Is Automated Journalistic Writing Less Biased? An Experimental Test of Auto-Written and Human-Written News Stories

Yanfang Wu

School of Communication, University of Miami, Miami, FL, USA

ABSTRACT

By administering an online experiment, this study examined how source and journalistic domains affect the perceived objectivity, message credibility, medium credibility, bias, and overall journalistic quality of news stories among an adult sample ($N = 370$) recruited using Amazon's Mechanical Turk (MTurk) service. Within the framework of the cognitive authority theory, the study found auto-written news stories were rated as more objective, credible (both message and medium credibility), and less biased. However, significant difference was found between a combined assessment condition (news stories with source and author information) and a message only assessment condition (news stories without source and author information) in the ratings of objectivity and credibility, but not bias. Moreover, significant differences were found in the objectivity and credibility ratings of auto-written and human-written news stories in the journalistic domains of politics, finance and sports news stories. In auto-written news stories, sports news stories were rated more objective and credible, while financial news stories were rated as more biased. In human-written stories, financial news stories were rated as more objective and credible. However, political news stories were rated as more biased among human-written news stories, and in cases where auto-written and human-written stories were combined.

Introduction

The incorporation of automation technologies into the workflow of newsrooms may represent positive or negative impacts on journalism (Barrat 2013; Pavlik 2016). Terms such as *automated*, *robot*, *algorithmic*, *machine-written*, or *computational journalism* are often used to describe when artificial intelligence, data analytics, and computational algorithms are applied to journalism (Anderson 2013a, 2013b; Stavelin 2014; Pavlik 2016; Carlson 2015; Graefe 2016; Thurman, Dörr, and Kunert 2017; Liu et al. 2017; Dörr 2016; Graefe et al. 2018). The aforementioned terms all come to the intersection of automated journalistic writing and journalism.

Automatic journalistic writing, which is involved with Human Computer Interaction (HCI) and Intelligence Augmentation (IA) in text generation, is becoming commonly

applied in news agencies (Linden et al. 2019). In the HCI process of journalistic writing, Natural Language Generation (NLG) is created based on digital structure data. However, when structured data is not available, IA advances automation where not all relevant information can be considered in a structured data set (Lavenda 2016).

It is commonly agreed that the growth of automated journalistic writing is critical to the future of news agencies and may affect how journalists create and how consumers perceive news (Graefe 2016; Marconi, Siegman, and Journalist 2017; Broussard 2015; Liu et al. 2017). For example, newsworthiness faces new constraints with digital news platforms; the "gut feeling" mode of journalistic judgment coexists with audience metrics for making news decisions (Lee, Lewis, and Powers 2014). Conflicting results have been found regarding the impact of automated journalistic writing on news production and consumption. Some scholars assert the implementation of automated journalistic writing in newsrooms could relieve journalists from routine tasks or repetitive part and allow them to focus on in-depth analysis and investigative journalism (Van Dalen 2012; Clerwall 2014; Marconi, Siegman, and Journalist 2017; Linden 2017b). Other researchers have concluded human journalists will likely be unable to compete with automated data collection and journalistic writing in terms of its speed, scale, and accuracy (Carlson 2015; Linden 2017b; Graefe 2016; Dörr 2016). Fear that automation will eventually eliminate newsroom jobs often dominates the debate (Van Dalen 2012; Carlson 2015). Wölker and Powell (2018) are in the middle ground. Their proposed method of combining human and automated writing in journalism has been practiced in newsrooms (Wu, Tandoc, and Salmon 2018; Young and Hermida 2015).

With advances in algorithms and the availability of big data, the integration of automated journalistic writing is on the rise (Dörr 2016; Thurman, Dörr, and Kunert 2017; Caswell and Dörr 2017). The disruption of automated journalism is likely to come, with the expectation that more than 90 percent of news will be written by machine reporters by 2025 (Levy 2012; Graefe et al. 2018). The growth of automated journalistic writing is mostly limited to data driven topics such as weather, routine sports, financial news (Caswell and Dörr 2017; Latar 2015; Waddell 2018; Kim, Oh, and Lee 2016), traffic reporting, earthquake warnings (Young and Hermida 2015), crime, business, politics, and disaster reporting (Caswell and Dörr 2017), although there is an incipient consensus that automated journalistic writing goes beyond that.

Journalism quality is measured through a series of clearly defined quality standards, including accuracy, balance, holding the government accountable, and separation of news from editorial and advertising. However, studies regarding the quality of automated journalistic writing are still thin and are limited to journalists' view about whether journalists think algorithms are less biased (Lazer et al. 2014; Latar 2015). Alternative quality journalism standards such as whether the journalistic writing is clear, coherent, comprehensive, concise, and well-written were not fully investigated. Graefe et al.'s (2018) conclusion that automated journalistic writing was highly readable was rendered less persuasive as the source was disclosed – written by a human reporter or algorithms – in the experiment. The experimental results showed that falsely declaring a machine-written article as written by a human journalist substantially increases the ratings of readability. Moreover, although an article's topic, its purpose, and consumers' expectations are regarded as factors that allow automated algorithms to augment or supplant journalists, how different types of news topics – like finance, sports or politics – and authorship –

human-written or machine-written – may affect readers' perception of quality of journalism is largely unknown (Graefe et al. 2018). Furthermore, whether message credibility – trustworthiness and expertise, will be more salient in the evaluation of news quality is uncertain. With the flood of information on the internet, readers have insufficient bandwidth for both the source and the message information. Due to redistribution by multiple sources or outlets, the original source is often unknown to consumers (Appelman and Sundar 2016). Thus, it is valuable to investigate how readers perceive the quality of auto-written journalism compared to human-written journalism when the confounding effects of source are excluded. In particular, comparing the message and medium credibility and objectivity, and how readers' expectation of quality vary among stories of different journalistic domains.

Literature Review

Automated Journalistic Writing (Auto-Written)

Natural language generation (NLG) technology – the automatic creation of text from digital structure data – is increasingly being used in automated journalistic writing (Caswell and Dörr 2017; Thurman, Dörr, and Kunert 2017; Reiter and Dale 1997). Motivated by the urgent need to increase the productivity of journalistic content, the use of auto-written content began with short texts that are generally indistinguishable from text written by human writers (Örnebring and Conill 2016; Graefe et al. 2018). In digital journalism, automation mostly refers to AI-enabled processes that include the availability of structured data and algorithms capable of drawing inferences (Lewis, Guzman, and Schmidt 2019). Auto-written text has been applied in journalism since the 1960s on weather forecasts (Linden 2017a; Glahn 1970), and since the 1990s on sports, medical and financial reports (Dörr 2016). However, automatic journalistic writing has not been implemented in more sophisticated and long form journalism such as event driven storytelling. (Dörr 2016; Caswell and Dörr 2017). It is obvious that not all journalistic work can be automated. New-writing technologies can't perform all the same tasks of human communicators. Simple template NLG systems are only useful in highly limited settings. However, it deserves to explore new opportunities by NLG in the media industry (Linden et al. 2019; Lewis, Guzman, and Schmidt 2019).

Auto-written journalism differs from human-written journalism in that it lacks the ability to recognize, create, organize or communicate stories (Thurman, Dörr, and Kunert 2017; Melin et al. 2018). Automatic writing is unable to appropriately use the nuances of language such as humor, sarcasm, and metaphors, which are valuable in more sophisticated storytelling (Deuze 2005a; Lewis and Westlund 2015; Thurman, Dörr, and Kunert 2017; Latar 2015; Dörr 2016). In consequence, auto-written stories tend to sound technical, boring, and less readable (Latar 2015). On the contrary, human reporters can produce stories that contain multi-sourced, interrogated, opinionated, and contextualized journalism with nuance and creativity (Thurman, Dörr, and Kunert 2017; Latar 2015; Marconi, Siegman, and Journalist 2017). Although auto-written and human-written journalism are not identical, auto-written stories are considered close enough to be a viable alternative (Lewis, Guzman, and Schmidt 2019). Despite its disadvantages, some scholars are optimistic about automated journalistic

writing in more complex stories. For example, Caswell and Dörr (2017) believe that complicated, event-driven stories can be written by automated extraction of structured event data from text, though this has not yet been perfected.

Economic pressures on the media industry may cause more news organizations to seek new solutions to the challenges imposed by the digital revolution (Marconi, Siegman, and Journalist 2017). More news organizations have started to practice automated journalistic writing because it can augment core journalistic skills such as accuracy and speed (Dörr 2016; Latar 2015; Cleary and Cochie 2011; Carlson 2018). Previous studies have found that the applications of automated journalistic writing may limit transaction costs for journalistic text production, thereby increasing speed, expanding coverage and leading to efficiency gains (Latzer et al. 2016; Marconi, Siegman, and Journalist 2017; Reeves 2016).

Objectivity

As a form of communication, journalism seeks to provide the public with information about issues that are fact-based (Weaver and Willnat 2012; Quandt 2000) or objective (Neuberger 1996), and avoid any authorial opinion or point of view (Thomson, White, and Kitley 2008). Detachment, nonpartisanship, facticity, and balance are major components of objectivity (Mindich 2000). Considered the marker of professional news and inscribed within standard journalistic conventions, objectivity has been used to evaluate the quality of journalism since the 1920s (Streckfuss 1990, 973; Shudson 1978). In particular, many believe that automated journalistic writing can greatly improve objectivity by reducing bias and eliminating human error caused by emotional motivations, fatigue and discouragement. Specifically, automated journalistic writing may undercut the subjective intentions in story selection and ordering (Sundar 2008; Ellul 1964; Reeves 2016; Graefe et al. 2018). Algorithms connect to a larger discourse that "contrasts inherent human subjectivity with the automated objectivity of computer programs" (Carlson 2018, 10). Studies have shown that news selection algorithms present a technological solution to what has long been considered a professional problem – lack of objectivity, and is potentially outstripping the profession's existing practices (Carlson 2017; Barnhurst and Nerone 2002). Algorithmic authority (Carlson 2015) and algorithmic objectivity (Gillespie 2014) have expanded professional ideologies – objectivity as journalism's "occupational ideology" toward journalistic quality (Deuze 2005b; Carlson 2018).

Opponents argue that algorithms could produce unexpected un-objective results if they rely on biased data based upon wrong or incomplete facts or make assumptions based on conscious or unconscious biases (Diakopoulos 2015; Latar 2015). Some researchers found that the development of new algorithms may place new pressures on journalistic judgement. In particular, the presence of bias in the algorithms may affect the objectivity of automated journalistic writing (Carlson 2015; Gillespie 2014; Carlson 2018; Toraman and Can 2015). How to evaluate the quality of automated journalistic writing without knowing how programmers developed the algorithm to generate auto-written news stories remains unclear. Savvy programmers attempting to game the system may even spread false stories (Carlson 2018). In response, Graefe and co-authors (2018) proposed to analyze how consumers perceive the quality of auto-written journalism in

relation to human-written journalism. Based on the scholarship in automatic journalistic writing, and Graefe and co-authors' proposal, the following research question was asked:

RQ1: How do readers assess automatic journalistic writing comparable to human writing?

Credibility

Credibility has been utilized to evaluate whether news is a valid reflection of reality (Newhagen and Nass 1989), and can be defined as believability – trustworthiness and expertise (Fogg and Tseng 1999). Starting from Hovland and Weiss' (1951) inquiries of source credibility in acceptance of persuasive messages, credibility research has been a primary concern in the communication research field for decades. Message credibility (Austin and Dong 1994), and medium or channel credibility (Bucy 2003; Johnson and Kaye 1998; Kiousis 2001) are regarded as important components of credibility. Newhagen and Nass (1989, 3) defined credibility as "the degree to which an individual judges his or her perceptions to be a valid reflection of reality" (p. 3).

Credibility is a multidimensional concept encompassing trustworthiness, believability, fairness, bias, incompleteness, accuracy, and social concerns. These are all important dimensions in news credibility judgement (Gaziano and McGrath 1986; Metzger et al. 2003; Meyer 1988; Meyer 1974; Sundar 1999). According to the cognitive theory, how an individual judges credibility could be based on many factors – channel, source, content, context, motives of the source, situation, and time, or any combination of these (Austin and Dong 1994; Delia 1976). However, many existing cognitive methods employed to differentiate the relative value of information may not be applied to the Internet. Due to the insufficient bandwidth for storing information (for both source and message credibility), the failure to adopt cognitive methods, and the cue cognitive heuristics effect when people process mediated messages, how people perceived the credibility of internet information was viewed as intentionally or unintentionally inaccurate, biased or misleading Also, known manipulation intent on the part of the source negatively impacts trustworthiness (Flanagin and Metzger 2000; Hernon 1995; Sundar 2008; Lang 2000).

Credibility should be measured on multiple dimensions as some dimensions are quite labile. For example, experimental results showed that a newspaper may lose its community affiliation credibility but still retain its straightforward believability (Meyer 1988). According to Gaziano and McGrath (1986), whether news organizations and journalists are concerned about the community's well-being and public interest are important indicators of credibility. However, the credibility construct may change under other circumstances. Among the threefold construct of credibility, the medium and/or source are important factors that affect how users perceive the credibility of a message (Wilson 1983). However, when medium and source are not revealed, users evaluate the message based on the perceived quality of the message (Newhagen and Nass 1989; Hovland and Weiss 1951; Flanagin and Metzger 2000; Slater and Rouner 1996). For example, Sundar's (1999) credibility index was characterized by "objectivity, fairness and bias". However, "accuracy" and "believability", which were frequently identified with the credibility construct in prior factor-analytic studies (for example, Meyer's (1988) credibility index), were not present in the factors. In addition, Appelman and Sundar (2016) proposed a new message credibility construct – accurate, authentic and believable – to evaluate how an individual judge the veracity of communication content on social media.

Former research findings showed that algorithmic judgement is transforming newsworthiness and news knowledge in journalism (Carlson 2018). However, auto-written and human-written journalism had conflicting effects on readers' perception of news qualities. For example, in Clerwall's (2014) and Van der Kaa and Krahmer's (2014) experimental studies, auto-written journalism received higher ratings on credibility, while human-written journalism scored higher on readability. News consumers valued auto-written journalism's trustworthiness, and were slightly negative about the expertise of a journalist writer. However, the studies often fail to distinguish between medium and message credibility, but tended to directly examine perceived content or message credibility.

Ethical challenges evolve with the rise of automatic journalistic writing at the organizational, professional, and social sphere. Transparency such as data source disclosure, disclosure of automation algorithms, as well as legal accountability on the output level become emerging ethical issues. Developing codify ethical guidelines for non-journalistic actors involved might be necessary as well. Although algorithmic journalism may also be able to relieve the ethical challenges in verifying multiple sources, reflection, or diligence (Dörr and Hollnbuchner 2017).

Despite researchers' endeavors, how readers evaluate quality of news differently in the internet age is far from being understood. The effect of automated journalistic writing on readers' perception of journalistic quality, in particular, objectivity – the marker of professional news, and credibility – whether news is a valid reflection of reality, without conflating it with assessments that are attributable to source and/or medium, is unknown. It is valuable to examine whether readers may value some indicators or dimensions of objectivity and credibility construct more than others with the rise of automated journalistic writing. Therefore, the second research question was proposed:

RQ2: How does the standard of journalistic quality evolve in the internet age?

Cognitive Authority Theory

Cognitive authority theory has been used for the interpretation of individuals' information processing on the internet (Fritch and Cromwell 2001; Rieh and Belkin 1998). Wilson (1983) explicitly discussed various external tests for recognizing cognitive authority, and notes that cognitive authority is related to credibility. Wilson divided cognitive authority theory into four dimensions: personal authority (author), institutional authority (publisher), textual type authority (document type), and intrinsic plausibility authority (content of text). A reader may go through all of the four dimensions to assess the content quality or credentials of a text. However, if there are certain dimensions missing, a reader cannot relate to their previous experience and previously established authority structures. They have to rely solely on the available dimension(s) to evaluate the quality or credibility of a text.

In this study, to distinguish institutional authority (publishing house or source) and personal authority (author) from intrinsic plausibility authority (content of text), experimental stories were created under two conditions: in one condition, the authorship and publishing house information were kept as it was in original news story. Participants relied on institutional authority (the publishing house of the article or source), personal authority (the author of the article), and intrinsic plausibility (textual authority) to assess the quality. However, in the second condition, authorship and publishing house information

was removed. Therefore, among the threefold construct that relates to credibility - the source and the medium were removed from the experimental stories (Haim and Graefe 2017). Participants had to use textual type authority based on the text itself, separate from its author or publisher, and intrinsic plausibility authority built on readers' own assessment of the plausibility of the content to evaluate the objectivity, credibility, and overall quality of the experimental news stories (Rieh 2010).

How will source, medium, and authorship affect readers' evaluation of the quality of automated journalistic writing? What are the most salient dimensions of medium and message objectivity and credibility in auto-written journalism? What are the most important indicators differentiating auto-written and human-written journalism in terms of medium and message objectivity and credibility? For example, bias—one of the most important dimensions of credibility that could be greatly reduced by automated algorithms—has not been verified with consistent results (Graefe 2016); Moreover, how does readers' expectation and perception of quality vary among stories in different journalistic domains such as finance, sports, and politics – when some are auto-written and some are human-written? In light of this gap in the research of auto-written journalism and human-written journalism, and based on Wilson's (1983) cognitive authority theory, the following hypothesized model were proposed. To secure the generalizability of this study, the hypotheses were tested in a general adult sample in the US Figure 1:

H1: An auto-written news story will be rated as more objective than a human-written news story in a message only assessment condition (2) than in a combined assessment condition (1).

H2: The message credibility of an auto-written news story will be rated higher than a human-written news story news story in a message only assessment condition (2) than in a combined assessment condition (1).

H3: The medium credibility of an auto-written news story will be rated higher than a human-written news story in a message only assessment condition (2) than in a combined assessment condition (1).

H4: A human-written news story will be rated as more biased than an auto-written news story in a message only condition (2) than in a combined assessment condition (1).
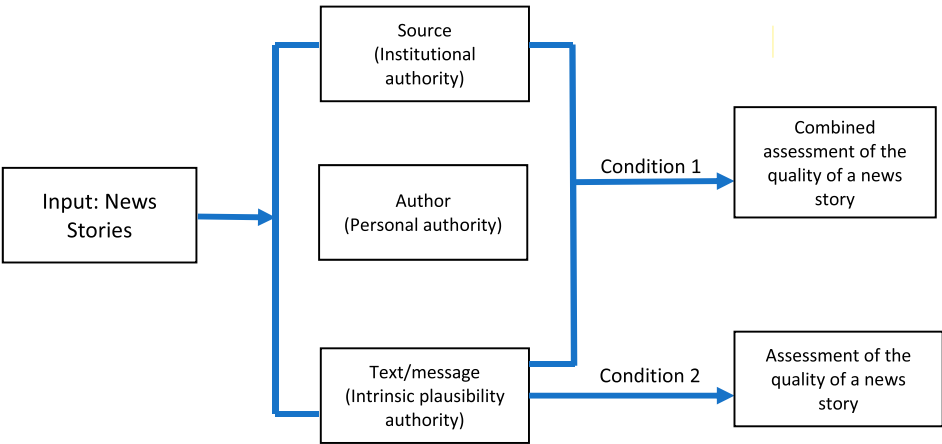

Figure 1. Hypothesized News Quality Assessment Model.

H5: A political news story will be rated as more biased than a sports news story. A sports news story will be rated as more biased than a finance news story.

H5a: An auto-written political news story will be rated as more biased than auto-written sports news story. An auto-written sports news story will be rated as more biased than an auto-written finance news story.

H5b: A human-written political news story will be rated as more biased than a human-written sports news story. A human-written sports news story will be rated as more biased than a human-written finance news story.

Methods

To uncover the differences in how readers evaluate auto-written and human-written journalism in relation to the descriptors of content or message objectivity and credibility, a 3 (journalistic domain: sports, finance, and weather; between subjects) x 2 (type of journalism: automated journalism, human journalism; within subjects) mixed, within-between subjects design online survey was executed via Qualtrics and Amazon Mechanical Turk - the crowdsourcing website. Altogether, there were six conditions in the experiment: a sports story by a machine writer, a sports story by a human writer, a finance story by a machine writer, a finance story by human writer, a political story by a machine writer, and a political story by a human writer.

The experimental survey was then executed in Amazon MTurk. In total, 370 adult participants from the United States were recruited on MTurk. Participant who completed the study on MTurk (and passed all attention filters) were compensated $0.50 for participating the study. The experiment took participants less than 20 min to complete. Participants provided implied consent before beginning the study. Participants' sex distribution was balanced with 185 (50%) females and 185 (50%) males in the MTurk crowdsourcing sample.

Experimental Design

A mixed-factorial experiment with type of stories as a between-subjects factor and type of journalist as a within-subjects factor was administered. Experimental stories were divided into two conditions – experimental stories with source and authorship information and experimental stories without source and authorship information. To manipulate the type of journalist (authorship or auto-written/human-written) conditions, one story generated by algorithms and one story written by a human reporter were selected from three different journalistic domains – sports, finance, and politics. All-together, six news stories were selected in the experimental test. The within subject design allowed each participant to read one story in each of the two conditions – auto-written and human-written news stories. The between subjects allowed each participant to read stories from one of the three journalistic domains – finance, sports, and politics.

Independent Variables (Manipulation)

Two independent variables were manipulated in the study: automation and journalistic domains. To manipulate the auto-written and human-written journalism conditions, recently published auto-written and human-written news stories from major national

news organizations were selected. To avoid potential confounding effects caused by the source on the perceived quality of news stories, news stories were selected from multiple news organizations — The Associated Press, Forbes, The Washington Post, and The New York Times. Auto-written news stories were generated by Automated Insights from Associated Press, or is powered by Heliograf, The Washington Post's artificial intelligence system. In the with source condition, the manipulation mimics how source is normally disclosed to readers in contemporary newsrooms. In the no source condition, the name of the news organization and authorship information were removed from the experimental stories. Since audiences are familiar with the mixture of auto-written and human-written stories in their daily lives, the manipulation of auto-written and human-written journalism is considered rational (Liu and Wei 2018).

To manipulate journalistic domains, the experimental stories chosen came from the domains of sports, finance and politics because previous studies found that automated journalistic writing had been widely applied in these fields. Additionally, to retain a controlled experiment, selected news stories in both auto-written and human-written journalistic conditions covered the same company, sports, and political issues or politicians, and each pair were of similar length. In particular, the auto-written news stories were either provided by the selected news organizations for this research study, or chosen from stories tagged as auto-written stories from selected news organizations' website.

Dependent Measures

Questionnaires were developed based on Donsbach and Klett's (1993) survey of journalists, Gaziano and McGrath's (1986), Meyer's (1988), Newhagen and Nass' (1989) credibility indexes, and Sundar's (2008, 1999) credibility and quality of news descriptions. Questions such as ratings of "good news reporting"- five dimensions including "no subjectivity," "fair representation," "fair skepticism," "hard facts," and "value judgment," - and then ratings of "objectivity", and the relationships between "good news reporting" and "objectivity" were asked; and a 10-point Likert-type scale were developed to gauge objectivity and credibility. Also, questions gauged fairness, bias, accuracy, fact, opinionatedness, separating fact from opinion, trustworthiness, as well as whether the news story was concerned with the public interest or the community's well-being, morality, or whether the story was written by well-trained reporters, at both the individual and organizational levels.

Pre-test

An initial pre-test of the stimuli materials, conducted with 36 undergraduate students from a large U.S. university in the Midwest, confirmed the comprehension difference of authorship. The manipulation check in the actual research resulted in significant differences (p = .001) between conditions.

Procedure

Participants were recruited from Amazon Mechanical Turk *(N = 370)* to take part in a study about "news story processing." After consenting to participate in the study, participants were assigned to two conditions – experimental stories with source and authorship

information *(N = 370)* and experimental stories without source and authorship information *(N = 370)*. Then participants were randomly assigned to the three groups: finance story group *(N = 238)*, sports story group *(N = 246)*, and politics story group *(N = 256)*. Participants were instructed to read one auto-written story and one human-written story, then respond to the same block of questionnaire for both.

Data Analysis

In order to determine how many factors to retain from the 42 items with scales, a parallel analysis was executed in SPSS. Comparing the eigenvalues between the random data matrices and the actual data, and using the eigenvalues that correspond to the desired percentile (95%) of the distribution of random data eigenvalues, components were retained as long as the eigenvalue from the actual data was greater than the eigenvalue from the random data.

According to Kline (1979), the subjects-to-variables (STV) ratio of 2 with a minimum of 100 subjects is acceptable. In this study, the STV ratio in the university student sample is 0.86 with 36 cases, which is not enough for a parallel analysis; however, four factors loaded together in the principle factor analysis. The sample from MTurk is 8.81 with 370 subjects, which satisfies Kline's (1979) standard to perform factor analysis. The analysis showed that in the MTurk sample, four factors needed to be retained in order to explain the relationships among the variables at a 95% confidence level.

In the next step, an exploratory factor analysis with four latent factors was conducted and a varimax rotation method was selected to determine the indicators of each latent factor in the MTurk sample. Factors that failed to load at least three indicators needed to construct a latent factor and were therefore excluded. A total of four factors were extracted from the MTurk data set.

This four-factor solution had a comparative fit index (CFI) of .83 in the measurement Structural Equational Modeling (SEM). Then the model was trimmed based on modification indices from the SPSS Amos recommendations until a CFI of .97 was reached, which is indicative of a good representation of the data (Teel and Verran 1991; Bentler 1990).

Based on the indicators of each factor, the researcher interpreted the factors as corresponding to the concepts (1) objectivity, (2) message credibility, (3) medium credibility, and (4) bias, which are explained in greater depth in the subsequent sections.

Objectivity (F1)

Three items were extracted to measure objectivity. Indicators included how strongly the participant agreed or disagreed with the following statements: (1) The story made it clear what is fact and what is opinion; (2) The story was objective; and (3) The story was a good news report in which the journalist's own political belief did not affect the presentation of the subject. A confirmatory factor analysis (principal axis factoring) using varimax as the rotation method was conducted for the four indicators of objectivity. The total variance after extraction was 6.581, accounting for 9.346% of variance. The reliability test showed that Cronbach's alpha for the three items of balance was .768.

Message Credibility (F2)

Five items were extracted to measure message credibility. Indicators included how strongly the participant agreed or disagreed with the following statements: (1) The story was a good news report that expressed views fairly from multiple sides about the event; (2) This story made it clear what is fact and what is opinion; (3) The story was objective; (4) The story was a good news report in which the journalist's own political beliefs did not affect the presentation of the subject; and (5) The story was ethical. A confirmatory factor analysis (principal axis factoring) using varimax as the rotation method was conducted for the eight indicators of message credibility. The total variance after extraction was 13.797, accounting for 32.851% of variance. The reliability test showed that Cronbach's alpha for the five items of message credibility was .903.

Medium Credibility (F3)

Four items were extracted to measure medium credibility. Indicators included how strongly the participant agreed or disagreed with the following statements: (1) This story was written by a well-trained writer; Based on the story, I think (2) the newspaper that carried the story has clear ethical principles; (3) the newspaper that carried the story is trustworthy; and (4) the news organization that carried the story does good journalism. A confirmatory factor analysis (principal axis factoring) using varimax as the rotation method was conducted for the four indicators of credibility. The total variance after extraction was 3.925, accounting for 9.346% of variance. The reliability test showed that Cronbach's alpha for the four items of medium credibility was .869.

Bias (F4)

Three items were extracted to measure bias. Indicators included how strongly the participant agreed or disagreed with the following statements: (1) This story seemed written from a biased viewpoint; (2) The story was predominantly opinionated; and (3) The journalist might not have had access to important facts that would change the story significantly. A confirmatory factor analysis (principal axis factoring) using varimax as the rotation method was conducted for the four indicators of bias. The total variance after extraction was 1.592, accounting for 3.791% of variance. The reliability test showed that Cronbach's alpha for the three items of bias was .809.

Results

A series of mixed factorial ANOVA tests were utilized to analyze the data and test hypotheses.

The test for hypothesis 1 showed that significant differences existed in the ratings of objectivity between auto-written news stories *(M = 20.08, SD = 4.11)* and human-written news stories *(M = 18.05, SD = 5.32), F (1, 736) = 43.123, p = .000 < .001*; The ratings of objectivity in a message only assessment condition (without source) was significantly higher *(M = 19.07, SD = 4.85)* than in a combined assessment condition (with source) *(M = 18.27, SD = 5.28), F (1, 736) = 4.58, p = .03 < .05.* H1 was supported.

The test for Hypothesis 2 showed that significant differences existed in the ratings of message credibility between auto-written and human-written news stories *(M = 32.87, SD = 6.87; M = 29.32, SD = 8.55), F (1, 736) = 36.072, p = .000 < .001.* Auto-written stories were rated more credible than human-written news stories; Moreover, the ratings of credibility in a message only assessment condition (without source) were significantly higher *(M = 31.83, SD = 7.73)* than in a combined assessment condition (with source) *(M = 30.59, SD = 8.07), F (1, 736) = 4.630, p = .03 < .05.* H2 was supported.

The test for Hypothesis 3 showed that significant difference existed between auto-written and human-written news stories in the ratings of medium credibility *(M = 19.85, SD = 4.43; M = 17.34, SD = 5.45), F (1, 736) = 43.123, p = .000 < .001.* The medium credibility of auto-written stories was rated higher than human-written news stories. Moreover, the ratings of medium credibility in a message only assessment condition (without source) was significantly higher *(M = 19.07, SD = 4.85)* than in a combined assessment condition (with source) *(M = 18.27, SD = 5.28), F (1, 736) = 4.580, p = .03 < .05.* H3 was supported by the data.

The test for Hypothesis 4 showed that significant differences existed between auto-written and human-written news stories in the ratings of bias *(M = 12.86, SD = 6.05; M = 15.51, SD = 5.21), F (1, 736) = 37.257, p = .000 < .001.* Auto-written stories were rated less biased than human-written news stories. Moreover, participants perceived news stories more biased in a message only assessment condition (without source) *(M = 14.49, SD = 5.95)* than in a combined assessment condition (with source) *(M = 13.97, SD = 5.68),* though the difference was not significant, *F (1, 736) = 1.53, p = .216 > .05.* H4 was not supported.

The test for Hypothesis 5 showed that among different journalistic domains, there was no significant difference in the ratings of bias, *F (2, 734) = .317, p = .728 > .05,* H5 was not supported. Political news stories were rated more biased *(N = 256, M = 14.41, SD = 5.79)* than finance news stories *(N = 238, M = 14.10, SD = 5.67);* Finance news stories were rated more biased than sports news stories *(N = 246, M = 14.03, SD = 6.05).* H5 was not supported by the data.

For auto-written news stories, finance stories *(N = 119, M = 13.50, SD = 5.78)* were rated more biased than sports news stories *(N = 123, M = 12.68, SD = 6.37);* sports news stories were rated more biased than politics news stories *(N = 128, M = 12.44, SD = 5.98).* H5a was not supported.

For human-written news stories, political news stories *(N = 128, M = 16.38, SD = 5.00)* were rated more biased than sports news stories *(N = 123, M = 15.38, SD = 5.40);* and sports news stories were rated more biased than finance news stories *(N = 119, M = 14.70, SD = 5.51).* H5b was supported.

Discussion

The study found that auto-written news stories were rated as significantly more objective than human-written news stories. This finding is in line with previous researchers' assumptions about the inherent objectivity of algorithms, the limits of humans' subjective "gut feeling" in the evaluation of newsworthiness and news inclusion, and the advantages of algorithms in overcoming inherent human biases and limitations (Carlson 2018; Toraman and Can 2015). Also, the results corroborated Cleary and coauthor's (2011)

conclusion that Natural Language Generation software could augment accuracy, Clerwall's (2014) result that text generated by algorithms is perceived as more informative, accurate, and objective, and Melin and coauthor's (2018) finding that auto-written content tend to be rated more accurate, trustworthy and objective. Additionally, these findings echoed Thurman and coauthors' (2017) conclusion on automaton and the increase of objectivity in news stories.

The reason why automatic journalistic writing was rated as more objective could be that readers favor stories distinguishing facts and opinions clearly, which NGL, the generation method used for the texts chosen in this study, was recognized as being able to generate stories of this nature in Melin and coauthor's (2018) conclusion. Further, Graefe and coauthor's (2018) recent studies concluded that algorithms, such as NLG, are more accurate on factual information. Moreover, choosing the right vocabulary that represents the information in numbers is major task for journalists. Word choice is always influenced by the journalist's personal interpretation, which may reduce a news story's objectivity. For example, Carlson (2018, 1764) believed journalists have inherent human subjectivity because they apply learned knowledge to professionally valid interpretations. In contrast, algorithms have the unthinking objectivity of computer programs, and are the apotheosis of journalistic knowledge production – objectivity. Furthermore, specialized algorithms have a narrow domain focus reducing the options for word choice, thereby increasing objectivity (McDonald 2010). Gillespie (2014) used the term "algorithmic objectivity" to describe the power of algorithms in strengthening objectivity. Additionally, the integration of automation and datafication in news reporting may increase objectivity. For example, web analytics are found to be useful tools in increasing the precision of journalists' news judgement (Wu, Tandoc, and Salmon 2018). Data driven journalism not only empowers journalists to report stories in new and creative ways (Anderson 2018), it is also believed to increase objectivity (Parasie and Dagiral 2013).

Another interesting finding is that auto-written stories were even perceived as more credible (both message and medium credibility) than human-written news stories. This finding aligns with Haim and Graefe's (2017), Van der Kaa and Krahmer's (2014), Clerwall's (2014), and Melin's (2018) conclusions that readers tend to perceive auto-written news as more credible than human-written stories. The finding also aligns with Wolker and Powell's (2018) claim that there are equal perceptions of credibility between auto-written and auto-and-human-combined-written content. Although auto-written algorithms lack the skills in using nuances of languages, human reporters may produce less credible news stories due to failing to express views, or distinguish facts from opinions clearly (Meyer, Marchionni, and Thorson 2010). However, an algorithm is viewed as a "credible knowledge logic" (Gillespie 2014, 191) because it is considered a force that could eliminate human biases or errors (Linden 2017a). Algorithms also create many more possibilities for detecting falsehood (such as bias, inaccuracy) automatically and verifying truth more effectively (Kaczmarek 2008). Stories developed by programmers from multi-sourced data can fulfill functions of professional journalism and may even align with more traditional journalistic standards (Parasie and Dagiral 2013; Dörr 2016). Furthermore, algorithms may perform better than human reporters in data verification, as algorithms restrict themselves to a specialized area with a very stipulated content (Perloff 1993; Hassan and Azmi 2018).

This experimental design distinguished the effect of source and journalists' authorship from text on readers' ratings of quality of journalism, which is one of the major

contributions of this study. This study further verified that readers consider automatic journalistic writing more objective when source and journalists' affiliation information were not disclosed. Moreover, the message and medium credibility of automated journalistic writing were rated significantly higher without source and authorship. These findings align with Graefe et al.'s (2018) results on the confounding effect of source on readers' ratings of credibility – the declaration of an article written by a journalist substantially increase the ratings of readability. The findings also showed a decline of trust in traditional media and human writers in producing good journalism, which was reflected in a recent survey by Knight Foundation and Gallup poll – a majority of surveyed Americans reported they had lost trust in American media due to inaccuracy, bias, fake news, alternative facts, a general lack of credibility, and reporters are "based on opinions or emotions" (Ingram 2018). Automated journalistic writing, however, showed strength in objectivity and credibility.

The confounding effect of source and authorship was also identified in the ratings of bias. In particular, whether journalists' political belief affects the presentation of the subject is one of the important indexes of message credibility in this study. Human-written stories were rated more biased than auto-written news stories. However, when source and authorship information were included, human-written stories were rated less biased. Although impartiality is recognized as one of the important journalistic ideologies, human journalists were tagged as partisan actors whose political beliefs affected their news decisions, although they define themselves as news professionals committed to a form of journalism marked by objectivity and neutrality (Patterson and Donsbagh 1996; DellaVigna and Kaplan 2007; Oosterhoff, Shook, and Ford 2018; Linden 2017a). Gaziano and McGrath (1986) identified that political bias was important factor affecting credibility in news reporting, particularly accuracy, fairness, responsibility, and role in criticism of government. With the proliferation of data, human-written news stories may contextualize the automated-generated content by using it as a multi-source from different perspectives (Carlson 2015; Thurman, Dörr, and Kunert 2017). This was described by Carlson (2018) as "a visible incursion of algorithmic judgment in the space of human editorial judgment" and Wolker and Powell (2018) as the well-rounded future of journalism. These applications are feasible to reduce human bias in journalism.

Participants in the message only assessment condition rated auto-written news stories as both more objective and more credible (both message credibility and medium credibility) than human-written news stories compared to participants in the combined assessment condition. According to the hypothesized news assessment model based on the cognitive authority theory, readers rely on textual authority (intrinsic plausibility) – whether the content is "accurate, authentic and believable" (Appelman and Sundar 2016, 73) – to execute evaluation when the affiliation of the news organization and the journalist's name were removed from the stories. However, when institutional authority and personal authority were combined with textual authority, readers combine textual authority with whether the source of message is "authoritative, reliable, reputable and trustworthy" (Appelman and Sundar 2016, 74) to make judgment. The results showed that this combined assessment process reduced news stories' perceived objectivity and credibility. In the internet age, readers may assess news stories with greater emphasis on textual authority due to the insufficient bandwidth for storing information (Sundar 2008). The findings from this study corroborates Wilson's (1983) cognitive authority

theory. When source – personal authority and institutional authority – are not revealed, readers have to use textual type authority (document type), and intrinsic plausibility authority (content of text) to evaluate the credibility of a news story. This may change how news quality is assessed in digital journalism. The findings further verified the result that auto-written news content is perceived as more objective and credible than human-written news stories.

Although automated journalistic writing received higher credibility ratings, it is also more likely to distribute fake news due to its dependence on data for source, processing and output. Ethical issues may arise when data is used without proper verification, transparency about the source and content in generation algorithms (Graefe 2016). In addition, whether the programmer, reporter, or editor will be responsible for the collection, analysis and presentation of data, and who should be held accountable for automated journalistic writing in which human reporters contributed to contextualize generated content, and algorithms acted as an intelligence augmenter, remain controversial topics in the field of automated journalistic writing.

Journalistic domains were found to affect readers' evaluations of objectivity, message credibility, and medium credibility, but not bias. Sports news stories were rated more objective and credible (both message and medium credibility) than finance and political news stories in auto-written news stories. Financial news stories were rated more objective and credible than sports and political news stories among human-written stories. Financial news stories were rated more biased than sports and political news stories among auto-written news stories. However, political news stories were rated as more biased than financial and sports news stories in human-written news stories. Political news stories were rated as more biased than sports and finance news stories when auto-written and human-written stories were combined. Multiple business motivations may result in journalistic domains having an effect on readers' assessment of news stories. First, the publishers' or new organizations' political stance was found to greatly affect that of its reporters'. For example, news stories of ABC, CBS, NBC and Fox were believed to exhibit political leanings. Subsequently, political bias may be more salient in story constructions or patterns of bias in news stories (Groeling 2008). Similarly, participants' political ideology may affect their views towards The Associated Press, The New York Times, and The Washington Post. Secondly, social identity – the sense of whether one belongs to a perceived in-group or out-group – may affect bias in sports news. Previous studies found that American sportscasters tended to report favorably on athletes from the United States and highlight them more frequently. In-group favoritism is more pronounced when their participants' team won (Bryant and Oliver 2009; Wann 2006). However, financial news stories, which are mostly free from political standing and social identity, were the least biased when compared to politics and sports stories in the human-written group. On the contrary, when neither political stance nor social identity plays a role, financial news stories were rated the most biased in the auto-written group.

Conclusion

By administering an online experiment, this study found that when personal authority and institutional authority information were missing, audiences follow the cognitive authority theory and rely on textual type and intrinsic plausibility authority to evaluate news stories.

Auto-written news stories were rated more objective, and credible (both message and medium credibility), and less biased. However, significant difference was found between a combined assessment condition (with source and author information) and a message only assessment condition (without source and author information) in the ratings of objectivity and credibility, but not bias. Moreover, significant differences existed in the ratings of objectivity, and credibility (both message and medium credibility) of auto-written and human-written news stories in the journalistic domains of politics, finance and sports. In auto-written news stories, sports stories were rated more objective and credible (both message and medium credibility) than finance and political stories. In human-written news stories, financial stories were rated more objective and credible than sports and political news stories. However, financial news stories were rated more biased than sports and political news stories among auto-written stories. Political news stories were rated as more biased than financial and sports news stories in human-written stories, and when auto-written and human-written stories were combined.

This study is not without limitations. The reliability of the MTurk crowd sourcing sample needs to be further tested, and the general public may have different impressions regarding political beliefs, nationalism, or in-group and out-group standpoints. Moreover, political views and brand bias toward The Associated Press, The New York Times, and The Washington Post might be other important factors affecting the experimental results. Furthermore, human factors such as human crafted templates in automated creation of texts could affect readers' perception of quality of journalism. Avenues for future research are plentiful. Future studies should endeavor to secure samples that better represent the overall population in terms of demographic information, education, and political ideology, and explore more stories from variety of news organizations. Also, only three types – politics, sports, and finance news stories – were included in the experiment. Future studies may include broader types of stories to test the difference between auto-written and human-written journalism.

Overall, using Wilson's (1983) cognitive authority theory, this study highlights how auto-written and human-written journalism differ in readers' perceptions of objectivity, credibility, and bias when personal authority and institutional authority were absent and when combined with textual authority. This contributes to uncovering of the uncertain effects of automated journalistic writing on readers' perception of the quality of journalism. The results and analysis of this study shed light on whether and how auto-written content contributes to journalistic objectivity, credibility (both message and medium credibility), and bias in a way that outstrips the profession's existing frameworks, and the overall qualities of a news story. The academic contributions of this experimental study are obvious: it initiated the application of cognitive authority theory into the quality evaluation of automated journalistic writing, distinguished the effect of textual authority from the combination of institutional authority, personal authority and textual authority, tested the effect of source and journalistic domains on message credibility and medium credibility between auto-written and human-written news stories. The academic findings can be empirically significant for the digital journalism industry in the long run.

Disclosure statement

ORCID

Yanfang Wu ⓘ http://orcid.org/0000-0003-2696-8047

References

Anderson, Christopher W. 2013a. "Towards a Sociology of Computational and Algorithmic Journalism." *New Media & Society* 15 (7): 1005–1021.

Anderson, Christopher W. 2013b. "What Aggregators do: Towards a Networked Concept of Journalistic Expertise in the Digital age." *Journalism* 14 (8): 1008–1023.

Anderson, Christopher William. 2018. *Apostles of Certainty: Data Journalism and the Politics of Doubt*. Oxford: Oxford University Press.

Appelman, Alyssa, and S. Shyam Sundar. 2016. "Measuring Message Credibility: Construction and Validation of an Exclusive Scale." *Journalism and Mass Communication Quarterly* 93 (1): 59–79.

Austin, Erica Weintraub, and Qingwen Dong. 1994. "Source v. Content Effects on Judgments of News Believability." *Journalism & Mass Communication Quarterly* 71 (4): 973–983.

Barnhurst, Kevin G., and John Nerone. 2002. *The Form of News: A History*. New York, NY: Guilford Press.

Barrat, James. 2013. *Our Final Invention: Artificial Intelligence and the end of the Human era*. London: Macmillan.

Bentler, Peter M. 1990. "Comparative fit Indexes in Structural Models." *Psychological Bulletin* 107 (2): 238.

Broussard, Meredith. 2015. "Artificial Intelligence for Investigative Reporting: Using an Expert System to Enhance Journalists' Ability to Discover Original Public Affairs Stories." *Digital Journalism* 3 (6): 814–831.

Bryant, Jennings, and Mary Beth Oliver. 2009. *Media Effects: Advances in Theory and Research*. Abingdon: Routledge.

Bucy, Erik P. 2003. "Media Credibility Reconsidered: Synergy Effects Between on-air and Online News." *Journalism & Mass Communication Quarterly* 80 (2): 247–264.

Carlson, Matt. 2015. "The Robotic Reporter: Automated Journalism and the Redefinition of Labor, Compositional Forms, and Journalistic Authority." *Digital Journalism* 3 (3): 416–431.

Carlson, Matt. 2017. *Journalistic Authority: Legitimating News in the Digital Era*. New York, NY: Columbia University Press.

Carlson, Matt. 2018. "Automating Judgment? Algorithmic Judgment, News Knowledge, and Journalistic Professionalism." *New Media & Society* 20 (5): 1755–1772.

Caswell, David, and Konstantin Dörr. 2017. "Automated Journalism 2.0: Event-Driven Narratives: From Simple Descriptions to Real Stories." *Journalism Practice* 12 (4): 477–496.

Cleary, Johanna, and Meredith Cochie. 2011. "Core Skill set Remains Same in Newspaper job ads." *Newspaper Research Journal* 32 (4): 68–82.

Clerwall, Christer. 2014. "Enter the Robot Journalist: Users' Perceptions of Automated Content." *Journalism Practice* 8 (5): 519–531.

Delia, Jesse G. 1976. "A Constructivist Analysis of the Concept of Credibility." *Quarterly Journal of Speech* 62 (4): 361–375.

DellaVigna, Stefano, and Ethan Kaplan. 2007. "The Fox News Effect: Media Bias and Voting." *The Quarterly Journal of Economics* 122 (3): 1187–1234.

Deuze, Mark. 2005a. "Towards Professional Participatory Storytelling in Journalism and Advertising." *First Monday* 10: 7.

Deuze, Mark. 2005b. "What is Journalism? Professional Identity and Ideology of Journalists Reconsidered." *Journalism* 6 (4): 442–464.

Diakopoulos, Nicholas. 2015. "Algorithmic Accountability: Journalistic Investigation of Computational Power Structures." *Digital Journalism* 3 (3): 398–415.

Donsbach, Wolfgang, and Bettina Klett. 1993. "Subjective Objectivity. How Journalists in Four Countries Define a key Term of Their Profession." *Gazette (Leiden, Netherlands)* 51 (1): 53–83.

Dörr, Konstantin Nicholas. 2016. "Mapping the Field of Algorithmic Journalism." *Digital Journalism* 4 (6): 700–722.

Dörr, Konstantin Nicholas, and Katharina Hollnbuchner. 2017. "Ethical Challenges of Algorithmic Journalism." *Digital Journalism* 5 (4): 404–419.

Ellul, Jacques. 1964. *The Technological Society*. New York: Knopf.

Flanagin, Andrew J., and Miriam J. Metzger. 2000. "Perceptions of Internet Information Credibility." *Journalism & Mass Communication Quarterly* 77 (3): 515–540.

Fogg, B. J., and Hsiang Tseng. 1999. "The Elements of Computer Credibility." Proceedings of the SIGCHI conference on Human Factors in Computing Systems.

Fritch, John W., and Robert L. Cromwell. 2001. "Evaluating Internet Resources: Identity, Affiliation, and Cognitive Authority in a Networked World." *Journal of the American Society for Information Science and Technology* 52 (6): 499–507.

Gaziano, Cecilie, and Kristin McGrath. 1986. "Measuring the Concept of Credibility." *Journalism Quarterly* 63 (3): 451–462.

Gillespie, Tarleton. 2014. "The Relevance of Algorithms." *Media Technologies: Essays on Communication, Materiality, and Society* 167–167.

Glahn, Harry R. 1970. "Computer-produced Worded Forecasts." *Bulletin of the American Meteorological Society* 51 (12): 1126–1131.

Graefe, Andreas. 2016. *Guide to Automated Journalism*. New York, NY: Tow Center for Digital Journalism, Columbia Journalism School.

Graefe, Andreas, Mario Haim, Bastian Haarmann, and Hans-Bernd Brosius. 2018. "Readers' Perception of Computer-Generated News: Credibility, Expertise, and Readability." *Journalism* 19 (5): 595–610.

Groeling, Tim. 2008. "Who's the Fairest of Them all? An Empirical Test for Partisan Bias on ABC, CBS, NBC, and Fox News." *Presidential Studies Quarterly* 38 (4): 631–657.

Haim, Mario, and Andreas Graefe. 2017. "Automated News: Better Than Expected?" *Digital Journalism* 5 (8): 1044–1059.

Hassan, Isyaku, and Mohd Nazri Latiff Azmi. 2018. "Islam-related News Credibility in Selected Nigerian and Malaysian Newspapers".

Hernon, Peter. 1995. "Disinformation and Misinformation Through the Internet: Findings of an Exploratory Study." *Government Information Quarterly* 12 (2): 133–139.

Hovland, Carl I, and Walter Weiss. 1951. "The Influence of Source Credibility on Communication Effectiveness." *Public Opinion Quarterly* 15 (4): 635–650.

Ingram, Mathew. 2018. "Most Americans say They Have Lost Trust in the Media." *Columbia Journalism Review*. https://www.cjr.org/the_media_today/trust-in-media-down.php.

Johnson, Thomas J, and Barbara K. Kaye. 1998. "Cruising is Believing?: Comparing Internet and Traditional Sources on Media Credibility Measures." *Journalism & Mass Communication Quarterly* 75 (2): 325–340.

Kaczmarek, Adam L. 2008. "Automatic Evaluation of Information Credibility in Semantic Web and Knowledge Grid." WEBIST (2).

Kim, Soomin, JongHwan Oh, and Joonhwan Lee. 2016. "Automated News Generation for TV Program Ratings." Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video.

Kiousis, Spiro. 2001. "Public Trust or Mistrust? Perceptions of Media Credibility in the Information age." *Mass Communication & Society* 4 (4): 381–403.

Kline, Paul. 1979. *Psychometrics and Psychology*. Cambridge, MA: Academic Press Inc.

Lang, Annie. 2000. "The Limited Capacity Model of Mediated Message Processing." *Journal of Communication* 50 (1): 46–70.

Latar, Noam Lemelshtrich. 2015. "The Robot Journalist in the Age of Social Physics: The End of Human Journalism?" In *The New World of Transitioned Media*, 65–80. Springer.

Latzer, M., K. Hollnbuchner, N. Just, and F. Saurwein. 2016. "The Economics of Algorithmic Selection on the Internet." In Handbook on the Economics of the Internet, edited by J. Bauer and M. Latzer, 395–425. Cheltenham: Edward Elgar.

Lavenda, D. 2016. "Artificial Intelligence vs. Intelligence Augmentation." *Network World (Online)*. http://search/. proquest. com/docview/1809109037.

Lazer, David, Ryan Kennedy, Gary King, and Alessandro Vespignani. 2014. "The Parable of Google Flu: Traps in big Data Analysis." *Science* 343 (6176): 1203–1205.

Lee, Angela M., Seth C. Lewis, and Matthew Powers. 2014. "Audience Clicks and News Placement: A Study of Time-Lagged Influence in Online Journalism." *Communication Research* 41 (4): 505–530.

Levy, Steven. 2012. "Can an Algorithm Write a better News Story than a Human Reporter? Wired." Accessed 0224. https://www.wired.com/2012/04/can-an-algorithm-write-a-better-news-story-than-a-human-reporter/.

Lewis, Seth C., Andrea L. Guzman, and Thomas R. Schmidt. 2019. "Automation, Journalism, and Human–Machine Communication: Rethinking Roles and Relationships of Humans and Machines in News." *Digital Journalism* 7 (4): 409–427.

Lewis, Seth C., and Oscar Westlund. 2015. "Big Data and Journalism: Epistemology, Expertise, Economics, and Ethics." *Digital Journalism* 3 (3): 447–466.

Linden, Carl-Gustav. 2017a. "Decades of Automation in the Newsroom: Why are There Still so Many Jobs in Journalism?" *Digital Journalism* 5 (2): 123–140.

Linden, Tommy Carl-Gustav. 2017b. "Algorithms for Journalism." *The Journal of Media Innovations* 4 (1): 60–76.

Linden, Tommy Carl-Gustav, Hanna Emilia Tuulonen, Asta Bäck, Nicholas Diakopoulos, Mark Granroth-Wilding, Lauri Haapanen, Leo Juhani Leppänen, Magnus Melin, Tom Arne Moring, and Myriam Douce Munezero. 2019. "News Automation: The Rewards, Risks and Realities of'machine Journalism'".

Liu, Xiaomo, Armineh Nourbakhsh, Quanzhi Li, Sameena Shah, Robert Martin, and John Duprey. 2017. "Reuters Tracer: Toward Automated News Production using Large Scale Social Media Data." 2017 IEEE International Conference on Big Data (Big Data).

Liu, Bingjie, and Lewen Wei. 2018. "Reading Machine-Written News: Effect of Machine Heuristic and Novelty on Hostile Media Perception." International Conference on Human-Computer Interaction.

Marconi, Francesco, Alex Siegman, and Machine Journalist. 2017. *The Future of Augmented Journalism: A Guide for Newsrooms in the age of Smart Machines*. New York: Associated Press.

McDonald, David D. 2010. "Natural Language Generation." *Handbook of Natural Language Processing* 2: 121–144.

Melin, Magnus, Asta Bäck, Caj Södergård, Myriam D. Munezero, Leo J. Leppänen, and Hannu Toivonen. 2018. "No Landslide for the Human Journalist-An Empirical Study of Computer-Generated Election News in Finland." *IEEE Access* 6: 43356–43367.

Metzger, Miriam J., Andrew J. Flanagin, Keren Eyal, Daisy R. Lemus, and Robert M. McCann. 2003. "Credibility for the 21st Century: Integrating Perspectives on Source, Message, and Media Credibility in the Contemporary Media Environment." *Communication Yearbook* 27: 293–336.

Meyer, Timothy J. 1974. "Media Credibility: The State of the Research." *Public Telecommunications Review* 2 (4): 48–52.

Meyer, Philip. 1988. "Defining and Measuring Credibility of Newspapers: Developing an Index." *Journalism & Mass Communication Quarterly* 65 (3): 567–574.

Meyer, Hans K., Doreen Marchionni, and Esther Thorson. 2010. "The Journalist Behind the News: Credibility of Straight, Collaborative, Opinionated, and Blogged "News"." *American Behavioral Scientist* 54 (2): 100–119.

Mindich, David T. Z. 2000. *Just the Facts: How" Objectivity" Came to Define American Journalism*. New York, NY: NYU Press.

Neuberger, Christoph. 1996. *Journalism as Problem Processing. Objectivity and Relevance in Public Communication*. UVK: Konstanz.

Newhagen, John, and Clifford Nass. 1989. "Differential Criteria for Evaluating Credibility of Newspapers and TV News." *Journalism Quarterly* 66 (2): 277.

Oosterhoff, Benjamin, Natalie J. Shook, and Cameron Ford. 2018. "Is That Disgust I see? Political Ideology and Biased Visual Attention." *Behavioural Brain Research* 336: 227–235.

Örnebring, H., and R. F. Conill. 2016. "Outsourcing Newswork." In *The SAGE Handbook of Digital Journalism*, edited by T. Witschge, C. Anderson, and D. Domingo, 207–221. London: Sage.

Parasie, Sylvain, and Eric Dagiral. 2013. "Data-driven Journalism and the Public Good:"Computer-Assisted-Reporters" and "Programmer-Journalists" in Chicago." *New Media & Society* 15 (6): 853–871.

Patterson, Thomas E., and Wolfgang Donsbagh. 1996. "News Decisions: Journalists as Partisan Actors." *Political Communication* 13 (4): 455–468.

Pavlik, John. 2016. "Cognitive Computing and Journalism: Implications of Algorithms, Artificial Intelligence and Data for the News Media and Society." *Brazilian Journal of Technology, Communication, and Cognitive Science* 4 (2): 1–14.

Perloff, Richard M. 1993. *The Dynamics of Persuasion: Communication and Attitudes in the 21st Century*. Abingdon: Routledge.

Quandt, Thorsten. 2000. "The End of Journalism? Online Com- Munication as a Challenge for Journalism Research." In *Theories of Journalism*, edited by Martin Lo ̈ffelholz, 483–509. Westdeutscher Verlag: Wiesbaden.

Reeves, Joshua. 2016. "Automatic for the People: The Automation of Communicative Labor." *Communication and Critical/Cultural Studies* 13 (2): 150–165.

Reiter, Ehud, and Robert Dale. 1997. "Building Applied Natural Language Generation Systems." *Natural Language Engineering* 3 (1): 57–87.

Rieh, Soo Young. 2010. "Credibility and Cognitive Authority of Information." *Encyclopedia of Library and Information Sciences* 1 (1): 1337–1344.

Rieh, Soo Young, and Nicholas J. Belkin. 1998. "Understanding Judgment of Information Quality and Cognitive Authority in the WWW." Proceedings of the 61st annual meeting of the american society for information science.

Shudson, Michael. 1978. "Discovering the News." *A Social History of American Newspapers*.

Slater, Michael D., and Donna Rouner. 1996. "How Message Evaluation and Source Attributes may Influence Credibility Assessment and Belief Change." *Journalism & Mass Communication Quarterly* 73 (4): 974–991.

Stavelin, Eirik. 2014. "Computational Journalism. When Journalism Meets Programming".

Streckfuss, Richard. 1990. "Objectivity in Journalism: A Search and a Reassessment." *Journalism Quarterly* 67 (4): 973–983.

Sundar, S. Shyam. 1999. "Exploring Receivers' Criteria for Perception of Print and Online News." *Journalism & Mass Communication Quarterly* 76 (2): 373–386.

Sundar, S. Shyam. 2008. "The MAIN Model: A Heuristic Approach to Understanding Technology Effects on Credibility." In *Digital Media, Youth, and Credibility*. Vol. 7., edited by Miriam J. Metzger and Andrew J. Flanagin, 73–100. The John D. and Catherine T. MacArthur Foundation Series on Digital Media and Learning. Cambridge, MA: The MIT Press.

Teel, Cynthia, and Joyce A. Verran. 1991. "Focus on Psychometrics. Factor Comparison Across Studies." *Research in Nursing & Health* 14 (1): 67–72.

Thomson, Elizabeth A., Peter R. R. White, and Philip Kitley. 2008. ""Objectivity" and "Hard News" Reporting Across Cultures: Comparing the News Report in English, French, Japanese and Indonesian Journalism." *Journalism Studies* 9 (2): 212–228.

Thurman, Neil, Konstantin Dörr, and Jessica Kunert. 2017. "When Reporters Get Hands-on with Robo-Writing: Professionals Consider Automated Journalism's Capabilities and Consequences." *Digital Journalism* 5 (10): 1240–1259.

Toraman, Cagri, and Fazli Can. 2015. "A Front-Page News-Selection Algorithm Based on Topic Modelling Using raw Text." *Journal of Information Science* 41 (5): 676–685.

Van Dalen, Arjen. 2012. "The Algorithms Behind the Headlines: How Machine-Written News Redefines the Core Skills of Human Journalists." *Journalism Practice* 6 (5–6): 648–658.

Van der Kaa, Hille, and Emiel Krahmer. 2014. "Journalist Versus News Consumer: The Perceived Credibility of Machine Written News." Proceedings of the Computation+ Journalism Conference, Columbia University, New York.

Waddell, T. Franklin. 2018. "A Robot Wrote This? How Perceived Machine Authorship Affects News Credibility." *Digital Journalism* 6 (2): 236–255.

Wann, D. 2006. "The Causes and Consequences of Sport Team Identification." In *Handbook of Sports and Media*, edited by A. Raney and J. Bryant, 331–352. New York: Routledge.

Weaver, David H., and Lars Willnat. 2012. *The Global Journalist in the 21st Century*. Abingdon: Routledge.

Wilson, Patrick. 1983. "Second-hand Knowledge: An inquiry into Cognitive Authority".

Wölker, Anja, and Thomas E. Powell. 2018. "Algorithms in the Newsroom? News Readers' Perceived Credibility and Selection of Automated Journalism." *Journalism* 1–18. doi:10.1177/1464884918757072.

Wu, Shangyuan, Edson C. Tandoc Jr, and Charles T. Salmon. 2018. "Journalism Reconfigured: Assessing Human–Machine Relations and the Autonomous Power of Automation in News Production." *Journalism Studies*. 1–18.

Young, Mary Lynn, and Alfred Hermida. 2015. "From Mr. and Mrs. Outlier to Central Tendencies: Computational Journalism and Crime Reporting at the Los Angeles Times." *Digital Journalism* 3 (3): 381–397.