Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Full length article

Intrusion of software robots into journalism: The public's and journalists' perceptions of news written by algorithms and human journalists

Jaemin Jung a, Haeyeop Song a,*, Youngju Kim b, Hyunsuk Im a, Sewook Oh b

a School of Management Engineering, KAIST, Seoul, South Korea
b Korea Press Foundation, Press Center Bldg., Taepyeongno 1-ga, Jung-gu, Seoul, 100-750, South Korea

ARTICLE INFO

Article history:
Received 11 July 2016
Received in revised form 6 February 2017
Accepted 6 February 2017
Available online 7 February 2017

Keywords:
Robot journalism
Journalist
Algorithms
Technology adoption
Ingroup-outgroup distinction

ABSTRACT

This study adopted a two (author: algorithm or journalist) by two (notification of author: real or inverse) between subject design to investigate how the public and journalists perceive the quality of algorithms-written articles compared with human journalist's work. Findings showed that both the public and journalists' evaluations were varied by the manipulation of author notification. That is, the public gave higher scores to the algorithm's work when it was notified as the real author, but they gave lower scores to the algorithm's work when the author was notified as a journalist. It confirmed the public's negative attitude toward journalists' credibility and craving for new information and communication technology (ICT) products/services in Korea. Based on journalists' resistance to change and innovation and the theory of prejudice, it was expected that journalists would be favorable to another journalist's work and unfavorable to an algorithm's work. However, contrary to the hypothetical expectation, journalists also gave higher scores to an algorithm's work and lower scores to a journalist's work. Implications relating to the intrusion of algorithm-written articles into journalism were discussed.

© 2017 Elsevier Ltd. All rights reserved.

1. Introduction

The rise of the Internet has changed not only how news is distributed and consumed but also how it is produced. The Internet has drawn more people and organizations into news production. Beyond journalists in traditional news companies, there are also professional bloggers and online news startups. Even the general public acts as news providers. A recent development in the realm of journalism is software-generated content. Journalism entered a new phase with the rise of computer-written or automatically produced news articles. This advanced technological development has led to a new type of journalism—robot journalism.

News companies have started to work with algorithms to operate and publish software-generated news articles. Templates are produced in journalistic but iterative processes that require painstaking manual work. Although the technology is still in an early market phase, automated journalism has arrived in newsrooms. For example, Forbes uses an artificial-intelligence platform provided by the technology company Narrative Science to generate automated news on corporate earnings and stock prices from live datasets and content harvested from previous articles (Gani & Haddou, 2014). AP partnered with Automated Insights to begin automating quarterly earnings reports and now publishes 3000 such financial stories every quarter (Miller, 2015). After an earthquake hit Los Angeles one morning, it took only 3 min for the LA Times to write and publish an article about it online. It was written by an algorithm (Neal, 2014).

Beyond the U.S. news media, companies in other regions introduced algorithm-written articles. Chinese social and gaming giant Tencent published its first business report written by an algorithm in September 2015. It was written in Chinese and completed in just 1 min by Dreamwriter, a Tencent-designed algorithm (He, 2015). The Financial News in Korea ran an article reporting stock market news written by IamFNBOT, an algorithm writer (Son, 2016).

Algorithms are used not only in news writing but also in news selection and editing. In the Guardian from the U.K., the algorithmic newspaper software selects the most popular Guardian articles and
assembles them into a weekly printed newspaper (Ellis, 2013). It is well known that search engines like Google also use algorithms to display news on its site.

As the technologies are intruding into the creation of news, scholars have started to examine the changing nature of journalism amid data abundance, computational exploration, and algorithmic emphasis with growing significance for the media industry and for journalism as practice and profession (Lewis, 2015; also see the special issue: “Journalism in an era of big data” of Digital Journalism, 2015). Among the data-oriented practices emerging in journalism, Carlson (2015) noted that none appear to be as potentially disruptive as automated journalism, insofar as it calls up concerns about the future of journalistic labor, news compositional forms, and the very foundation of journalistic authority. Meanwhile, automated content with no human intervention beyond the initial programming can make journalists free from handling basic works and afford time to focus on more investigative reporting. It could be an opportunity for media companies to reinvent news production system by generating news faster, at a larger scale, and with fewer errors. It also could be beneficial for audience pursuing more news and unbiased reporting.

In fact, journalists are not among the first to feel both the pressure and opportunity of automation. There have been periodic warnings in the last two centuries that automation and new technology were going to wipe our large number of middle class jobs (Autor, 2015). Such concerns have recently regained prominence. In their widely discussed book The Second Machine Age, Brynjolfsson and McAfee (2014) offered an unsettling picture of the likely effects of automation on employment. They asserted that there’s never been a better time to be a worker with special skills or the right education, because these people can use technology to create and capture value. However, there’s never been a worse time to be a worker with only ‘ordinary’ skills and abilities to offer, because computers, robots, and other digital technologies are acquiring these skills and abilities at an extraordinary rate (Brynjolfsson & McAfee, 2014, p. 11).

In this new era of automated journalism, to activate discussion on automated news creation and its implications for journalists, journalism, news industry, and audience, it is necessary to explore the public’s and journalists’ perceptions and evaluation of news written by algorithms. How does the public perceive and evaluate the quality of computer-written news articles? Can they tell the differences between articles written by a human journalist and algorithms? Journalists perceive themselves as a professional group differentiated with the public, because they have expertise and duty. Then, how do journalists’ perception and evaluations of algorithm-written articles compare with those from the public? This study compares the perceived quality of algorithm-written articles with the work by human journalists. To answer the questions, we sampled both general readers and journalists and conducted a series of experiments with human- and algorithm-written articles.

2. Literature review

The discourse around the use of computers and software to gather, distribute, and publish content has different kinds of labels (Clerwall, 2014). One term embracing it broadly is “computational journalism,” which is described as “the combination of algorithms, data, and knowledge from the social sciences to supplement the accountability function of journalism” (Hamilton & Turner, 2009, p. 2). Other terms are “robot journalism” (Clerwall, 2014; Dawson, 2010; Van Dalen, 2012) and “automated content” and “algorithmic news” (Anderson, 2013; Bunz, 2010; Levy, 2012).

Computational and algorithmic journalism might reshape the cultural practice of news creation due to the hybrid nature of newsroom sense-making technologies. Anderson (2013) argues that the human becomes partially obdurate and the material partly intentional. For this reason, there have been increasing attention to the impact of algorithm and data in journalism. Some of recent studies focused on the computational journalism using qualitative interviews and textual analysis. Young and Hermida (2015) found that computational thinking and techniques emerged in a (dis) continuous evolution of organizational norms, practices, content, identities, and technologies that interdependently led to new product. Carlson (2015) contended with the emergent practice of automated news content creation both in how it alters the working practices of journalists and how it affects larger understandings of what journalism is, and how it ought to operate.

Other studies examined the relationship between journalism and data. Anderson (2015) investigated historized relationship between journalism and big data, and then asserted we need to consider the material objects (whether interviews, documents, human observations, or other objects) that underlie journalistic processes. Parasie (2015) examined the question about to what extent would the processing of huge datasets allow journalists to produce new types of revelation based on the epistemological approach. Diakopoulos (2015) identified the algorithmic power as something worthy of scrutiny by computational journalists interested in accountability reporting. He tried to show how transparency might be used to effectively adhere to journalistic norms in the use of newsroom algorithms. Using algorithm as part of the journalistic process is not, by any means, a new phenomenon. However, the use of software to actually write a news story is a new advancement (Clerwall, 2014). Being rather new, less attention has been given to automatically generated articles and their implications for journalistic practice and audience response, especially in empirical studies.

With the advent of algorithms-written articles, scholars attempted to test the quality of the automatically produced article. The first study was conducted by Christer Clerwall (2014) in Sweden. He investigated how readers perceive software-generated content in relation to similar content written by a journalist. Although there were no statistically significant differences due to the limited sample size—46 respondents—the experiment revealed interesting and mixed results. While the software-generated content was perceived as more descriptive and boring, it was also considered to be more informative, objective, accurate, and trustworthy. In this study, the respondents were also asked to assess whether the text had been written by a journalist or by a computer. Of the 27 respondents who read the software-generated text, 10 thought a journalist wrote it and 17 thought it was also considered to be more informative, objective, accurate, and trustworthy. In this study, the respondents were also asked to assess whether the text had been written by a journalist or by a computer. Overall, journalistic content produced by an algorithm was not or was barely discernible from content written by a journalist in this experiment.

Another study explored the perceived credibility of algorithm-written news articles, searching specifically for differences and similarities between journalists and news consumers in the Netherlands (Van der Kaa & Krahmer, 2014). In this study, 168 native Dutch speakers and 64 Dutch journalists were asked to evaluate the perceived levels of the expertise and trustworthiness of news articles (sports and finance news) written by algorithms. Similar to the previous study (Clerwall, 2014), news consumers perceived the levels of trustworthiness and expertise of the algorithm and journalist equally.
The lack of differences between algorithm-generated content and articles written by human journalists (Clerwall, 2014; Van der Kaa & Krahmer, 2014) may be seen as an indicator that the software is doing a good job, or it may indicate that the journalist is doing a poor job—or perhaps both are doing a good (or poor) job.

The articles tested in the previous study (Van der Kaa & Krahmer, 2014), which investigate the quality of software-generated news, were written by algorithms, and only the author (computer or journalist) was manipulated. The manipulations, “this article is written by a computer” or “this article is written by a journalist,” were shown on every page. Findings showed that journalists perceived the trustworthiness of a journalist to be much higher than that of an algorithm. In other words, journalists gave a higher score to the article noted as “a work by journalist,” although it was the same content. On the other hand, consumers valued the perceived trustworthiness of a computer writer, although they were slightly negative about the expertise of a journalist. Again, they evaluated the same content and were only manipulated by the notification of the author (computer or journalist).

Previous attempts that explored the perceived quality of algorithm-written news articles yielded similar results. Both the public and journalists failed to recognize the differences between news written by an algorithm and a journalist. The quality evaluations of the articles were varied by manipulation of the author. However, they did not figure out why the respondents answered in a different way. This study assumes that the general public’s and journalists’ stereotypes and different attitudes toward algorithms and journalists might yield different perceptions of news written by algorithms and human journalists.

2.1. The public’s attitude toward journalists and technology

The image of journalists in popular culture is quite masculine. They are aggressive, self-reliant, tough, ambitious, cynical, cocky, and unsympathetic. Female journalists who are more compassionate, caring, maternal, and sympathetic have fought to overcome this contradiction throughout the 20th century and are still fighting the battle today (Saltzman, 2003). Most portrayals of working journalists highlighted many negative characteristics that branded the profession for decades (Evensen, 2000; Welford, 1998). Bridger (1997) analyzed 59 films containing portrayals of photojournalists and found the stereotypes framed in negative images. A number of movies such as “Network” and “Broadcast News” cynically depicted journalists as greedy, aggressive, and cold-hearted people who are manipulating the news-gathering process to achieve their professional success. Regardless of the real lives of journalists, they are stereotyped negatively.

Even in the context of South Korea, the place in which the current study is conducted, journalists are blamed as unreliable and unethical because of a recent event. There was a Sewol ferry tragedy on April 16, 2014. The South Korean ferry capsized while carrying 476 people, who were mostly high school students. In all, more than 300 passengers and crewmembers died in the disaster. The country’s mainstream media critically failed in their jobs right in the midst of the disaster. Public criticism has been mounting against the Korean media because they dismissed the truth, published provocative stories, and parroted the government’s announcements without checking the facts and questioning them. A growing number of people have called reporters “giregi”—a mix of Korean words meaning reporter and trash (Saltzman, 2003). Most portrayals of working journalists highlighted many negative characteristics that branded the profession for decades (Evensen, 2000; Welford, 1998). Bridger (1997) analyzed 59 films containing portrayals of photojournalists and found the stereotypes framed in negative images. A number of movies such as “Network” and “Broadcast News” cynically depicted journalists as greedy, aggressive, and cold-hearted people who are manipulating the news-gathering process to achieve their professional success. Regardless of the real lives of journalists, they are stereotyped negatively.

Even in the context of South Korea, the place in which the current study is conducted, journalists are blamed as unreliable and unethical because of a recent event. There was a Sewol ferry tragedy on April 16, 2014. The South Korean ferry capsized while carrying 476 people, who were mostly high school students. In all, more than 300 passengers and crewmembers died in the disaster. The country’s mainstream media critically failed in their jobs right in the midst of the disaster. Public criticism has been mounting against the Korean media because they dismissed the truth, published provocative stories, and parroted the government’s announcements without checking the facts and questioning them. A growing number of people have called reporters “giregi”—a mix of Korean words meaning reporter and trash (Saltzman, 2003). Most portrayals of working journalists highlighted many negative characteristics that branded the profession for decades (Evensen, 2000; Welford, 1998). Bridger (1997) analyzed 59 films containing portrayals of photojournalists and found the stereotypes framed in negative images. A number of movies such as “Network” and “Broadcast News” cynically depicted journalists as greedy, aggressive, and cold-hearted people who are manipulating the news-gathering process to achieve their professional success. Regardless of the real lives of journalists, they are stereotyped negatively.

Even in the context of South Korea, the place in which the current study is conducted, journalists are blamed as unreliable and unethical because of a recent event. There was a Sewol ferry tragedy on April 16, 2014. The South Korean ferry capsized while carrying 476 people, who were mostly high school students. In all, more than 300 passengers and crewmembers died in the disaster. The country’s mainstream media critically failed in their jobs right in the midst of the disaster. Public criticism has been mounting against the Korean media because they dismissed the truth, published provocative stories, and parroted the government’s announcements without checking the facts and questioning them. A growing number of people have called reporters “giregi”—a mix of Korean words meaning reporter and trash (Saltzman, 2003). Most portrayals of working journalists highlighted many negative characteristics that branded the profession for decades (Evensen, 2000; Welford, 1998). Bridger (1997) analyzed 59 films containing portrayals of photojournalists and found the stereotypes framed in negative images. A number of movies such as “Network” and “Broadcast News” cynically depicted journalists as greedy, aggressive, and cold-hearted people who are manipulating the news-gathering process to achieve their professional success. Regardless of the real lives of journalists, they are stereotyped negatively.

Even in the context of South Korea, the place in which the current study is conducted, journalists are blamed as unreliable and unethical because of a recent event. There was a Sewol ferry tragedy on April 16, 2014. The South Korean ferry capsized while carrying 476 people, who were mostly high school students. In all, more than 300 passengers and crewmembers died in the disaster. The country’s mainstream media critically failed in their jobs right in the midst of the disaster. Public criticism has been mounting against the Korean media because they dismissed the truth, published provocative stories, and parroted the government’s announcements without checking the facts and questioning them. A growing number of people have called reporters “giregi”—a mix of Korean words meaning reporter and trash (Saltzman, 2003). Most portrayals of working journalists highlighted many negative characteristics that branded the profession for decades (Evensen, 2000; Welford, 1998). Bridger (1997) analyzed 59 films containing portrayals of photojournalists and found the stereotypes framed in negative images. A number of movies such as “Network” and “Broadcast News” cynically depicted journalists as greedy, aggressive, and cold-hearted people who are manipulating the news-gathering process to achieve their professional success. Regardless of the real lives of journalists, they are stereotyped negatively.

Even in the context of South Korea, the place in which the current study is conducted, journalists are blamed as unreliable and unethical because of a recent event. There was a Sewol ferry tragedy on April 16, 2014. The South Korean ferry capsized while carrying 476 people, who were mostly high school students. In all, more than 300 passengers and crewmembers died in the disaster. The country’s mainstream media critically failed in their jobs right in the midst of the disaster. Public criticism has been mounting against the Korean media because they dismissed the truth, published provocative stories, and parroted the government’s announcements without checking the facts and questioning them. A growing number of people have called reporters “giregi”—a mix of Korean words meaning reporter and trash (Saltzman, 2003). Most portrayals of working journalists highlighted many negative characteristics that branded the profession for decades (Evensen, 2000; Welford, 1998). Bridger (1997) analyzed 59 films containing portrayals of photojournalists and found the stereotypes framed in negative images. A number of movies such as “Network” and “Broadcast News” cynically depicted journalists as greedy, aggressive, and cold-hearted people who are manipulating the news-gathering process to achieve their professional success. Regardless of the real lives of journalists, they are stereotyped negatively.

Even in the context of South Korea, the place in which the current study is conducted, journalists are blamed as unreliable and unethical because of a recent event. There was a Sewol ferry tragedy on April 16, 2014. The South Korean ferry capsized while carrying 476 people, who were mostly high school students. In all, more than 300 passengers and crewmembers died in the disaster. The country’s mainstream media critically failed in their jobs right in the midst of the disaster. Public criticism has been mounting against the Korean media because they dismissed the truth, published provocative stories, and parroted the government’s announcements without checking the facts and questioning them. A growing number of people have called reporters “giregi”—a mix of Korean words meaning reporter and trash (Saltzman, 2003). Most portrayals of working journalists highlighted many negative characteristics that branded the profession for decades (Evensen, 2000; Welford, 1998). Bridger (1997) analyzed 59 films containing portrayals of photojournalists and found the stereotypes framed in negative images. A number of movies such as “Network” and “Broadcast News” cynically depicted journalists as greedy, aggressive, and cold-hearted people who are manipulating the news-gathering process to achieve their professional success. Regardless of the real lives of journalists, they are stereotyped negatively.

Even in the context of South Korea, the place in which the current study is conducted, journalists are blamed as unreliable and unethical because of a recent event. There was a Sewol ferry tragedy on April 16, 2014. The South Korean ferry capsized while carrying 476 people, who were mostly high school students. In all, more than 300 passengers and crewmembers died in the disaster. The country’s mainstream media critically failed in their jobs right in the midst of the disaster. Public criticism has been mounting against the Korean media because they dismissed the truth, published provocative stories, and parroted the government’s announcements without checking the facts and questioning them. A growing number of people have called reporters “giregi”—a mix of Korean words meaning reporter and trash (Saltzman, 2003). Most portrayals of working journalists highlighted many negative characteristics that branded the profession for decades (Evensen, 2000; Welford, 1998). Bridger (1997) analyzed 59 films containing portrayals of photojournalists and found the stereotypes framed in negative images. A number of movies such as “Network” and “Broadcast News” cynically depicted journalists as greedy, aggressive, and cold-hearted people who are manipulating the news-gathering process to achieve their professional success. Regardless of the real lives of journalists, they are stereotyped negatively.
whose authority stems from discursive source operating both inside and outside the professional sphere. However, one obvious thing is professional claims serve to draw boundary line between those on the “inside” and “outside of the profession.

Historical cases reveal how the journalists have been drawn the boundaries between the outsider and the journalistic profession to sustain their hegemony. In the mid-1970s, Computerized Typesetting System (CTS) and Computer Assisted Reporting (CAR) as new technology have deeply affected journalism as labor. CTS have substituted the metal type in newspaper printing and CAR have substituted handwriting in reporting. The computer reduced the labor intensity of journalists in editing and reporting (Howard, 2014). These technologies also reduced the number of employees in the newsroom. For example, the ‘operator’ as technician of the metal type has been disappeared by computer. But, disappeared occupational group in newsroom such as operator did not belong to journalistic profession. These groups were assistant of journalists. Of course, journalists had been compelled to adopt computer as a new reporting and editing device. However, journalists thought of computer as assistant of them. Because they thought of themselves as professionals that are not substituted essentially at that time (Deuze, 2005).

The rise of Internet and other forms of participatory media can serve as an example of how change invariance to new technological developments (e.g., Lowrey & Mackay, 2008; Singer, 2003, 2005; Steensen, 2011). In the mid-1990s, journalists defended their hegemony over journalistic work by stressing the reliability of the information they provide, owing to their professional skills. Thereby they intended to “strengthen journalists’ claims over the tasks of gathering news and ensuring accuracy” (Lowrey, 2006, p. 494). Ornebring (2013) interviewed with professional journalists and showed that the journalists turn to traditional claims based on expertise and duty when professional journalists want to demarcate the boundary between them and citizen journalism.

Journalists face another challenge from machine-written news—the robot journalist. Again they are forced to position themselves in relation to an external threat, which potentially challenges their jurisdiction over their main tasks. Van Dalen (2012) studied the response of the journalistic community to automatic content creation by analyzing 68 newspaper articles and blog posts covering Statsheet, an automated sports news creation site, during 2010. Contrary to previous studies that described the journalistic community as conservative and defensive, journalists who wrote about robot journalists did not reject the new development. Van Dalen (2012) interpreted this finding for three different reasons. First, the journalists who wrote about robot journalists can be seen as early adopters who are more open to change than the journalistic community at large. Second, because automated content creation is mostly applied to sports journalism rather than domains like politics or international affairs, journalists might be more generous. Finally, the journalists who wrote about automated-content analysis are observant of a trend that still remains largely abstract and does not influence their own work directly.

Cognitive psychologists have found that people are significantly affected by instability in their environments, tend to fear the unknown, and strive to the status quo (Brocker, Grover, Reed, Dewitt, & O’Malley, 1987; Wanberg & Banas, 2000). Introduction of a new system in media organizations has commonly faced resistance from journalists and yielded negative results. Giles (1995) argued in his book on newsroom management that resistance is a rational response to change because change inevitably involves losses in routines, relationships, and traditions. Although there were positive or neutral responses on the newsroom convergence experiment (Dupagne & Garrison, 2006; Erdal, 2009), a number of innovation projects, including newsroom restructuring, adoption of computer systems in newsrooms, experimenting with multimedia and social networking, and partnering with other media, have been fleeting and peripheral to the core operation (Liu, 2006; Lowrey, 2005, 2009, 2011; Aviles, Leon, Sanders, & Harrison, 2004; Daniels & Hollifield, 2002; Domingo, 2008; Gade, 2004; Garcia, 2008; Russial, 2009; Sheffer & Schultz, 2009; Sterling, 2008).

With journalists’ negative reactions to innovation, it is necessary to pay attention to journalists’ occupational homogeneity as an ingroup. Journalists might draw a boundary between the outsider and the journalistic profession (Van Dalen, 2012). Contemporary research on intergroup relations, prejudice, and discrimination appears to accept, at least implicitly, the idea that ingroup favoritism and outgroup negativity are reciprocally related. The idea can be derived from a number of theoretical assumptions. In an environment of scarcity, individuals needed to band together in groups to compete with other groups for survival (Sumner, 1906). Another approach to ingroup-outgroup distinction can be explained by the zero-sum perspective in which attachment and positive affect toward an ingroup is achieved through distance and negative affect toward corresponding outgroups. In fact, it is a sense universally true that “we” are more peaceful, trustworthy, friendly, and honest than “they.” This is reinforced by a general preference for the familiar over the unfamiliar (Brewer, 1995). Following zero-sum perspective, journalists would perceive algorithms replace their job. Otherwise, journalist would perceive the algorithms’ works as complementary.

Thus, we might expect that journalists have more positive attitudes toward ingroup members’ (human journalists’) work and negative attitudes toward outgroup members’ (algorithms’) work. Following the same logic, we can expect that journalists will evaluate a journalist’s work lower when the author is noted as an algorithm.

H2. Regardless of the real author, journalists are more likely to be favorable to a journalist’s article and less favorable to an algorithm’s work. That is, journalists’ evaluations will be different by manipulation of author notification.

3. Methods

3.1. Selection of stimulus

This study selected baseball game articles generated by software that developed by a university research team in Korea. Automated journalism requires high-quality data in structured and machine-readable formats. For this reason, automation works particularly well in domains such as finance, sports, or weather, where data providers make sure that the underlying data are accurate and reliable (Graefe, 2016). Other topics in different sections such as politics, economics, or culture might have more relevance with societal level. However, this study has no choice but to use baseball game articles due to the limited technology in Korea at this moment.

The articles used in the experiments should be representative for both algorithm-written content and human journalists’ work. We found that algorithm-generated articles are very similar in format and content. We compared the similarity of 10 articles by asking 20 people to indicate the quality (well written, clear, professional, and credible) with a 5-point Likert scale. There were no statistical differences among the articles, and we selected 1 of the 10 articles randomly.

In comparison to algorithm-written work, content written by human journalists varies in quality and quantity. Thus, more attention should be paid in selecting representative human journalists’ articles. First, we selected 10 articles reporting the same-
day ball game result. These articles were written by human journalists working for different news companies and were shortened in order to match the length of the one written by the algorithm. Then, we asked 20 respondents to select one article that seemed to be most similar with the algorithm’s work already chosen. Additionally, these 20 respondents evaluated the quality of the 10 articles based on the four components (well written, clear, professional, and credible). The results of both tests for selecting a human article yielded the same results. Eight respondents out of 20 selected one article as the most similar to the example written by an algorithm, and the mean score of that article ($M = 3.23$) was almost equal to the mean of the selected algorithm article ($M = 3.22$).

3.2. Pretest: comparing an algorithm’s and journalist’s work without author notification

We conducted a pretest on 201 people (103 male and 98 female) that were recruited from a national online panel to validate the stimuli used for the quality evaluation ($M_{\text{age}} = 39.6, SD = 11.24$). A professional survey company administered the survey. Respondents were given a small reward for completing the questionnaire. The research firm, Hankuk Research in Seoul Korea, emailed a link containing a questionnaire to respondents listed on the panel of the research firm.

Participants read one of two articles (written by an algorithm or journalist) without any notification about the author. After reading the article, participants indicated the extent to which they agree on a 5-point scale (1 = not at all to 5 = extremely) regarding the four components of article quality (well written, clear, professional, and credible). The four items were grouped as one factor, and the Cronbach alpha for internal reliability was 0.81. As we expected, there was no statistical difference between the two groups (algorithm: $M = 3.21$; journalist: $M = 3.31$; $F = 1.269, p = 0.261$). Results of the pretest confirmed that the general public failed to discriminate the quality of the algorithm’s and human journalist’s articles when the author of the articles was not disclosed. Thus, it validated the appropriateness of the stimuli.

4. Study 1: evaluation by the general public

Study 1 was designed to test hypothesis 1: the general public’s preference for an algorithm-written article over a human journalist’s work. When the author is noted correctly, the public would give higher scores on the algorithm’s work. However, the evaluation of the algorithm’s work would be lower when the author was noted as the journalist. Participants were presented with an article written by an algorithm or human journalist and asked to evaluate the quality of the article.

4.1. Participants and procedure

This study consisted of 400 individuals (200 male and 200 female) recruited from a national online panel who were asked to read an article and then answer a series of questions ($M_{\text{age}} = 39.3, SD = 11.60$). Subjects recruited from a professional research company, Hankuk Research in Seoul Korea. This study adopted a two (author of the article: algorithm or journalist) by two (notification of author: real or inverse) between subject design. The survey was distributed to 6295 people randomly. Among them, 1934 people opened the link and 1335 tried to begin the survey. The survey process was continued until a total of 400 completed questionnaires had been gathered. Note that though the remaining respondents also completed the questionnaires, they were not included in the final analysis because of late responses and budget limitations. The 400 participants were randomly assigned to one of four conditions until each quota is over. The manipulations, “this article is written by an algorithm” or “this article is written by a journalist,” were shown on the instructions. One group evaluated an article written by an algorithm and was instructed as such. The second group evaluated an article written by an algorithm but instructed as reverse. That is, although the article was written by an algorithm, they were told it was a human journalist’s work. The third group evaluated an article written by a human journalist and was instructed as such. The participants of the fourth group were assigned to the article written by a human journalist but were instructed as reverse. After reading the assigned article, each respondent answered the quality of the article (well written, clear, professional, and credible) with a 5-point Likert scale (1 = not at all to 5 = extremely).

4.2. Results

We tested our prediction using an analysis of variance (ANOVA) with the author of the article and notification of the author as the factors. As shown in Fig. 1, the author of the article by notification of the author interaction was significant ($F = 6.301, p < 0.05$). The public gave a higher score ($M = 3.39$) to the algorithm’s work when it was notified as the real author, but they gave a lower score ($M = 3.21$) to the algorithm’s work when the author was noted as a journalist. Although the article was written by an algorithm, evaluation was lower when the instructions stated it was a journalist’s work. Meanwhile, the public gave a lower score ($M = 3.31$) to the journalist’s article when the author was noted as a journalist, but they gave a higher score ($M = 3.41$) to the journalist’s work when the author was noted as an algorithm. Thus, regardless of the real author, the public is likely to be more favorable to an algorithm’s work and less favorable to a human journalist’s article. $H_1$ was supported.

Based on the first study, we figured out that the public gave higher evaluation to an algorithm’s work when it was noted as the real author, but they gave lower scores to an algorithm’s work when the author was notified as the reverse. On the other hand, the public gave lower evaluation to the human journalist’s article when the author was noted as a journalist, but they gave higher scores to the journalist’s work when the author was noted as the reverse. Thus, whether quality evaluation by the public on an algorithm’s work is higher than a journalist’s work depends on the manipulation of author notification.

5. Study 2: evaluation of journalists

Study 2 was designed to test hypothesis 2: journalists’ preferences for a human journalist’s article over an algorithm-written
article. When the author is disclosed truthfully, journalists gave higher scores to a journalist’s work. However, the evaluation would be lower when the author was noted as an algorithm. Participants were presented with an article written by an algorithm or human journalist and asked to evaluate the quality of the article.

5.1. Participants and procedure

There were 164 journalists (111 male and 53 female) recruited via snowballing who were given a link to an article at random order and were asked to read it and then answer a series of questions evaluating the quality of the article (Mage = 38.5, SD = 7.12). Snowball sampling is a nonprobability approach to sampling design and inference in hard-to-reach, or equivalently, hidden population. Standard statistical sampling method requires a list of population members from which the sample can be drawn. Although snowball sampling for journalist hinders the generalization of the results, this study adopted the snowball sampling method because it is difficult to gather journalists’ samples systematically. This study selected the first participants among the list of journalists, which participated the previous research in Korea Press Foundation. Study 2 also adopted a two (author of the article: algorithm or journalist) by two (notification of author: real or inverse) between subject design. Participants were assigned to one of four conditions identical to Study 1. Each cell had about 40 participants. The manipulations, “this article is written by an algorithm” or “this article is written by a journalist,” were shown on the instructions. Two groups were aware of the correct author of each article, and the other two groups read the article with wrong information about the author of the article. After reading the assigned article, each respondent ranked the quality of the article (well written, clear, professional, and credible) on a 5-point Likert scale (1 = not at all to 5 = extremely).

5.2. Results

To test H2, journalists’ preferences for a human journalist’s work over an algorithm’s work, ANOVA was conducted. As shown in Fig. 2, interaction between the author of the article and the notification of author was significant (F = 8.984, p < 0.01). When the author of the article was noted as the real author, journalists reading an algorithm’s article rated higher (M = 3.23), but they gave a lower score (M = 2.85) to an algorithm’s work when the author was noted as a journalist. Journalists gave a lower score (M = 3.13) to a journalist’s work when the author was noted as a journalist, but they gave a higher score (M = 3.33) to a journalist’s work when the author was notified as an algorithm. In Study 2, we figured out that journalists gave a higher evaluation to an algorithm’s work when it was noted as the real author, but they gave lower scores to an algorithm’s work when the author was noted as a journalist. On the other hand, journalists gave a lower evaluation on a human journalist’s article when the author was noted as a journalist, but they gave higher scores to a journalist’s work when the author was noted as an algorithm. Thus, whether quality evaluation by journalists on an algorithm’s work is higher than on a journalist’s work depends on the manipulation of author notification. It should be noted, however, that the direction of interaction was against our hypothetical expectation. That is, regardless of the real author, journalists are more likely to be favorable to an algorithm’s work and less favorable to a human journalist’s article. Thus, H2 was not supported. Detailed explanation is provided in the next section.

6. Conclusion and discussion

The current study compared the public’s and journalists’ perceptions of algorithm-written article with that of a human journalist’s article and investigated the effect of manipulation by author notification with a series of experiments. Consistent to previous studies testing the quality of algorithm-written articles in different country settings, this study confirmed that the public failed to discern an algorithm-written article from a human journalist’s article and gave undifferentiated evaluation on quality without author notification context. It proved that although algorithm-written articles are in their initial stages in Korea, the quality of algorithm-written articles is quite comparable to that of human journalists.

Regarding the public’s evaluation of article quality, we found interaction effect of author by true/reverse notification of author. That is, the public gave higher scores on the algorithm-written article when the author was noted as an algorithm. However, they gave lower scores on the algorithm’s article when the author was noted as a human journalist by manipulation. On the other hand, the public gave lower scores on a journalist’s article when the author was noted truthfully. They gave higher scores on a journalist’s article when the author was noted as an algorithm. It was the expected results based on the public’s negative attitude toward journalists’ credibility and craving for new ICT products/services in Korea.

The results of the journalists’ evaluations were against expectations. According to the previous theory of prejudice, journalists would give higher scores to ingroup members (i.e., a journalist’s work) and lower scores to outgroup members (i.e., an article written by an algorithm). However, journalists gave scores in the opposite way in this study. That is, journalists also gave higher scores to the algorithm-written article when the author was noted as an algorithm but gave lower scores to the algorithm’s work when the author was noted as a human journalist. As the public did, journalists also gave lower scores to the journalist’s article when the author was noted truthfully. However, they gave higher scores to the journalist’s article when the author was noted as an algorithm.

Presumably, it might be interpreted as the following reasons. Allport (1954) recognized that attachment to one’s ingroup does not necessarily require hostility toward the outgroup. Brewer (1999) empirically confirmed Allport’s contention that ingroup favoritism and outgroup prejudice are separable phenomena and that the origin of identification and attachment to ingroups is independent of intergroup conflict. Following Allport’s (1954) and Brewer’s (1999) alternative views, it is assumed that although human journalists in this experiment perceive a journalist who wrote the tested article as a member of the ingroup, the journalists’ ingroup identification did not necessarily lead to negative evaluation of the algorithm.
Moreover, in the absence of realistic conflict between human journalists and algorithm writers, neither strong ingroup attachment nor outgroup hostility was effective. In fact, since this study only tested a sports article, journalists in general did not perceive strong ingroup attachment to the human article. Since algorithm-written articles are published in sports and stock market reporting areas in a limited way in Korea, journalists did not seem to show any hostility to algorithm-written news yet.

Different from the zero-sum perspective between ingroup attachment and outgroup distance, journalists do not see the rise of the algorithm writer as a replacement that leads to a zero-sum game. Rather, journalists seem to expect algorithms would work as complementary in journalism by helping human journalists to have more time to focus on more investigative reporting. According to previous literatures, journalists framed an algorithm as a tool to facilitate systematic coverage and freeing journalists to add more depth, context, and the human touch, as well as possibly decrease costs (Young & Hermida, 2015). Van Dalen (2012) reached a similar conclusion about the impact of algorithmic journalism. Professional claims obviously serve to draw boundary lines between those on the inside and outside of the profession. Similar to CTS case, journalists seem to consider an algorithm as assistant and believe the algorithm writer as an unbiased reporter. Tow Center for Digital Journalism (2016) adds that journalists see the advantages of an algorithm as an unbiased reporter. Tow Center for Digital Journalism suggests other potentials for algorithms in journalism (Graefe, 2016). Algorithms are able to generate news faster, at a larger scale, and potentially with fewer errors than human journalists. Algorithms can use the same data to tell stories in multiple languages and from different angles, thus personalizing them to an individual reader’s preferences. Algorithms also have the potential to generate news on demand by creating stories in response to users’ questions about the data.

There are also concerns about this new phenomenon. Algorithms rely on data and assumptions, both of which are subject to biases and errors. As a result, algorithms could produce outcomes that were unexpected, unintended, and contained errors (Graefe, 2016). It might lead to the ethical quandary. If factual errors are discovered, does the software take responsibility? There are other ethical questions such as transparency and copyright. Although most journalists expect the robots will work as complementary to rather than replacing their jobs, automated content creation can be serious competition and a threat to the job security of journalists performing basic routine tasks (Van Dalen, 2012). In fact, changes in news organizations have been due to coercion from above and to weak-tie connectivity with the organization’s market and readers (Lowrey, 2011). To accommodate for an algorithm-written article article and develop high-quality journalism with the help of algorithms, more communication and accurate division of work roles within media organizations and understanding readers’ needs are necessary.

7. Limitations and suggestions

Despite its contributions to the literature on the perceptions of algorithm-written articles, this study has several limitations. First, the algorithm was developed to create sports articles due to the limited technology at this moment, so this study only tested the perceived quality of sports articles. Other news in different sections such as politics, economics, or culture might yield different results. Future research should take into account other content created by algorithms.

Second, this study sample consisted only of South Koreans. South Korea has highly developed broadband and mobile infrastructures and features a congested media-consumption environment. Nevertheless, perception and adoption of algorithm-written articles might differ across countries, and attitudes toward journalists and new ICT services could vary. Moreover, South Korea is a country with a collectivist cultural orientation. It would be valuable to investigate possible differences with other countries that are more individualistically inclined. Thus, the generalizability of our results is limited and requires caution.

Third, snowball sampling for journalists hinders the generalization of the results. Although acknowledging the difficulty in gathering journalists’ samples systematically, the diverse background and personal work roles might have influence on the perception of algorithm-written articles. Future research should pay more attention in sampling journalists.

Lastly, the articles used may not be very representative for either algorithm-written articles or content written by humans. To validate the representativeness of the tested articles, we followed a thorough sample selection process and performed a pretest with 201 samples. Nonetheless, we indicate that the experiment comprised one article from each category, making the risk of generalization.

References

Allport, G. W. (1954). The nature of prejudice. Cambridge, MA: Addison-Wesley.
