

Chinese Automated Journalism: A Comparison Between Expectations and Perceived Quality

CHENYAN JIA
University of Texas at Austin, USA

To explore the emerging practice of automated journalism in China, this study examined readers' expectations and actual perceptions of Chinese automated news through two experiments. Results showed that readers' actual perceptions of human-written news did not meet their expectations, but readers' actual perceptions of automated news exceeded expectations. When participants read both human-written and automated news, perceptions of human-written news were significantly higher for readability and expertise. When participants read either human-written or automated news, significant differences only existed for expertise. In both experiments, participants had similar expectations and perceptions of credibility between human-written and automated news. These results bear implications to the field in that the difference in credibility between automated and human-written news was indiscernible.

Keywords: Chinese automated journalism, credibility, expertise, expectation-confirmation theory, readability

The processes of media production and consumption have become automated and algorithmically dictated worldwide (Mager, 2012; Napoli, 2014). Major news organizations such as *The Associated Press* (United States), *The Los Angeles Times* (United States), *Reuters* (UK), and *The Financial News* (Korea) have adopted automation technology to produce news stories. In 2015, Tencent, one of the largest Chinese Internet companies, published the first piece of Chinese automated news written by an algorithm named Dreamwriter (Jung, Song, Kim, Im, & Oh, 2017). Although the original intention of automation was to free up human journalists from routine work by assisting journalists in analyzing a large amount of data (Graefe, 2016; Thurman, Dörr, & Kunert, 2017; Waddell, 2018), the readability, expertise, and credibility of automated news could be obstacles to its development because of the working principles of this new technological affordance. Compared with human-written news, algorithms have limitations in commentary, opinion, and investigative reports because automated news is currently restricted to repetitive and standardized narratives (Lokot & Diakopoulos, 2016).

Many scholars have articulated their concerns about the limitations of automated journalism and conducted empirical studies on the perceived quality of automated news in Western countries (e.g., Carlson, 2017; Clerwall, 2014; Waddell, 2018, 2019). Few studies, however, have examined readers' perceptions of

Chenyan Jia: chenyanjia@utexas.edu
Date submitted: 2019-09-07

Copyright © 2020 (Chenyan Jia). Licensed under the Creative Commons Attribution Non-commercial No Derivatives (by-nc-nd). Available at <http://ijoc.org>.

automated news written in Chinese. It is essential to study automated journalism in the context of China for several reasons. First, automated journalism is increasingly popular in China. During the 2016 Olympics, *Toutiao.com*, an online news portal in China, published more than 450 automated sports news stories that were read by nearly one million readers (Yin, 2017). During the 2018 FIFA World Cup, an artificial intelligence (AI) product named MAGIC developed by the *Xinhua News Agency* generated 37,581 pieces of automated news content and obtained approximately 120 million online page views (Fu, Shang, & Zhang, 2019). As of 2019, at least nine Chinese institutions (Tencent, Alibaba, Baidu, *Xinhua News Agency*, *Southern Metropolis Daily*, *Toutiao.com*, *China Business Network*, *People's Daily*, and *China Science Daily*) have developed journalism-oriented automation products, with the number continuing to rise. Second, China has become the largest Internet market in the world since 2008 (CIA, 2016). Even so, the Chinese Internet market remains underexplored and, to some extent, has been isolated from the global news market (Chyi, Lee, Tennant, 2019). Third, Chinese is the most spoken language in the world (Noack & Gamio, 2015). Although earlier studies have examined how journalists and news consumers perceive the quality of automated journalism written in English (Clerwall, 2014), German (Haim & Graefe, 2017; Graefe, Haim, Haarmann, & Brosius, 2018), and Dutch (Van der Kaa & Kraemer, 2014), few experimental studies have examined readers' responses to automated journalism written in Chinese language.

One pioneering work on Chinese automated news suggested people took cultural backgrounds into consideration when consuming automated content (Zheng, Zhong, & Yang, 2018). This comparative study tested how algorithms authorship cues affect Chinese and American news audiences' perceptions of automated news by manipulating the source (human or algorithms) of the content (Zheng et al., 2018). In the real world, however, the content generated by algorithms and written by human journalists is distinct in many aspects. Scholars suggested that studies that focused exclusively on the effects of source attribution may yield different results from studies that examined actual content produced by machines as opposed to humans (Graefe et al., 2018; Waddle, 2018). With that said, using stimuli generated by two entities is of importance. Therefore, this study explored the following overarching questions by using real stories either written by human journalists or by algorithms: What are Chinese readers' prior expectations and actual perceptions of automated journalism? Do they perceive automated news as readable, professional, or credible as human-written news?

Drawing on expectation-confirmation theory (ECT), the present study seeks to analyze readers' expectations and actual perceptions of automated news. Previous studies have shown that the perceived quality of human-written and automated news was significantly different when using within-subjects design, but not significantly different when using a between-subject design (Haim & Graefe, 2017). To ensure this cognitive perception process to be less affected by the study design, two experimental studies (Study 1: $n = 125$; Study 2: $n = 308$) were conducted to investigate the difference between the perceptions of automated and human-written news. Participants in Study 1 were randomly assigned to read either human-written or automated news. Participants in Study 2 were asked to read both human-written and automated news.

Literature Review

Automated Journalism

In recent years, algorithms have been reshaping the media sphere (Napoli, 2014; Thurman, 2011) and becoming a trending subject among journalism and communication researchers (Dörr, 2016). Automated journalism refers to a new format of news production that uses algorithms to automatically generate news content with little human intervention (Carlson, 2015; Haim & Graefe, 2017; LeCompte, 2015; Napoli, 2014; Zheng et al., 2018). Automated news or machine-written news is defined as the journalistic content created by algorithms (Thurman, Dörr, & Kunert, 2017; Waddell, 2019). An algorithm is a sequence of instructions designed by humans to tell a computer what to do (Domingos, 2015). Algorithms can create personalized content in multiple languages from different journalistic perspectives (Jung et al., 2017). When clean and structured data are available, algorithms can produce content more quickly and with potentially fewer errors than human journalists can (Graefe, 2016). This study avoided popular but misleading conceptualizations like “robot journalists” (Lokot & Diakopoulos, 2016) to refrain from calling attention to misleading depiction of automated journalism. In this work, the conceptual definition of Chinese automated news is news written in Chinese and generated by algorithms.

Algorithms have reshaped media production process by optimizing news content (Hagar & Diakopoulos, 2019). Early solutions for generating automated news were built on simple source codes designed to fill in the blanks of prewritten templates by pulling data from databases. With an increasing demand for automated content, news organizations have started using more sophisticated technologies to generate more readable material (Graefe, 2016). Likewise, Chinese automated journalism is no longer restricted to text, and in recent years has been expanding to encompass autogenerated audio and video, using cutting-edge algorithms like natural language generation technology, paraphrase generation, and text summarization.

Earlier studies related to the field of computational linguistics showed that even though the deep linguistic processing for Mandarin Chinese has been initialized in the past few years, the performance of this new technology on Chinese language is still far from satisfactory because of the challenges associated with in Chinese deep parsing (Sun, Chen, Wan & Liu, 2019). Given that developers are unlikely to fully disclose their algorithms, Diakopoulos (2015) proposed a promising but complicated approach called reverse engineering, which aims at decoding algorithms’ sets of rules by assessing the outcome. An alternative way for social scientists is to examine audiences’ receptivity to this new technological affordance to test the quality of automated news (Sundar, Waddell, & Jung, 2016). Adding to the literature, this work aims to examine Chinese readers’ expectations and the perceived quality of automated news.

Expectation-Confirmation Theory

ECT is a cognitive theory that provides a theoretical framework to examine consumer satisfaction, postpurchase behavior, and the adoption of new technologies (Bhattacharjee, 2001; Lin, Wu & Tsai, 2005). The theory was originally applied to marketing research by comparing the difference between preconsumption expectations and product performances (Oliver, 1980, 1993). Oliver and DeSarbo (1988) described ECT as the process in which consumers reached repurchase intentions. ECT posits that

disconfirmation between expectations and perceived perceptions has the most immediate influence on satisfaction (Oliver, 1993). If the actual perception confirms or exceeds the expectation, consumers are likelier to be satisfied and thus likelier to repurchase a product. Therefore, ECT assumes that satisfied consumers are likelier to repurchase the product than dissatisfied users (Haim & Graefe, 2017). There are two main attributes in this expectation-confirmation link, which are prior expectation and actual perception (or the perceived performance). Prior expectation is consumers' anticipation toward the product before purchase or usage behavior, which usually derives from personal needs, advertisements, or chatter (Oliver, 2015). Actual perception refers to consumers' evaluation of the product from their actual use experience.

ECT originally built on theories of attitude evaluation and cognitive dissonance (Festinger, 1957; Heider, 1946; Newcomb, 1959), and was later adopted for research in other social science fields. For news consumption literature, researchers defined reader perceptions as the evaluation of news stories (Sundar, 1999). As is true for other consumer purchase behaviors, news consumers also accept or reject the news information because of perceived news quality (Chung, 2017; Slater & Rouner, 1996). If readers' expectations of automated news are not met, negative expectancy violations can occur and undermine readers' news perceptions (Burgoon, 2015; Waddell, 2018). It is noteworthy that not all readers hold the same expectations of automated content, as previous studies reveal that recall of automation entities from past media use experience can significantly shape audiences' perceptions of automated news (Sundar et al., 2016).

Haim and Graefe (2017) first transferred ECT to the context of automated news in a study exploring the difference between people's prior expectations and actual perceptions of automated news and human-written news by conducting two experiments. Both within-subject and between-subject experiments showed that participants had higher expectations of human-written news compared with automated news about the articles' journalistic quality and readability, which confirmed Van der Kaa and Kraemer's (2014) finding that participants were positively surprised by the high quality of automated news. Following previous studies, this work adopted ECT to study readers' expectations and actual perceptions of Chinese automated news and sought to examine the following research question:

RQ: What is the difference between Chinese readers' perceptions of automated news and those of human-written news?

ECT expects that when people are positively surprised by an article's quality, they will assign higher quality ratings; when people are negatively surprised by an article's quality, they will assign lower quality ratings. Haim and Graefe (2017) argued that a news consumer's expectation toward an automated article was low for a lack of knowledge about the underlying technology. Because of the results of previous studies, this study predicted that the same pattern would be evident in the Chinese context, hypothesizing that:

H1: Chinese readers' prior expectations of human-written news are higher than those for automated journalism.

H2a: For Chinese human-written news, readers' actual perceptions are lower than their prior expectations.

H2b: For Chinese automated news, readers' actual perceptions are higher than their prior expectations.

Readability, Expertise, and Credibility

Previous studies have concentrated on quality assessments of automated journalism (Wölker & Powell, 2018). Some work used perceived news quality as a broad measurement (Jung et al., 2017; Zheng et al., 2018). Some studies examined more specific measurements such as the credibility of automated news by adapting the approach taken in Sundar's (1999) research into automated journalism research (e.g., Clerwall, 2014; Haim & Graefe, 2017). Others also adopt journalistic expertise and readability of automated news as measurements (Graefe et al., 2018; Van der Kaa & Kraemer, 2014).

Readability is a linguistic measure referring to the ease with which one can read content, which relies largely on the author's ability to simplify words, grammar, and sentences (Dubay, 2004; Fountas & Pinnell, 1999; Meyer, 2004). Human journalists often try to enhance the readability of news texts either by reducing vocabulary and grammatical burden (Dalecki, Lasorsa, & Lewis, 2009) or by building intrigue into a story. Journalistic expertise includes knowledge, experience, and competence (Van der Kaa & Kraemer, 2014). Graefe et al. (2018) did an exploratory factor analysis, which found that journalistic expertise includes measures of whether content is coherent, concise, comprehensive, and descriptive. The definition of credibility dates to the post-World War II era (Hovland, Janis, & Kelley, 1953). Credibility has been defined as perceptions of the receivers (Metzger, Flanagin, Eyal, Lemus, & McCann, 2003), including multiple factors such as trustworthiness, expertise (Chung, Nam, & Stefanone, 2012), and objectivity of the story (Sundar, 1999).

Previous studies on automated journalism have examined all three central tenets (readability, expertise, and credibility) in determining overall news quality. So far, the majority of related research demonstrated that participants expected more from human-written news in terms of readability (Haim & Graefe, 2017) and preferred human-written news for readability (Graefe et al., 2018; Haim & Graefe, 2017; Van der Kaa & Kraemer, 2014). Therefore, this study hypothesized that:

H3a: In terms of Chinese readers' actual perceptions, the readability of human-written news is higher than the readability of automated news.

Evidence is unclear as to whether the expertise of automated news is higher or lower compared with human-written news. Van der Kaa and Kraemer (2014) analyzed how journalists and news consumers perceived the expertise of news articles if they were ignorant of the article's source. The results of this study indicated that news consumers' ratings of expertise did not differ between human-written and automated news (Van der Kaa & Kraemer, 2014). Graefe et al. (2018) randomly assigned 986 participants to one of four conditions, in which they received either a human-written or an automated article with either correctly or wrongly declared authorship. Graefe et al., (2018) found that differences in perceptions of expertise tended to be small. This study predicted that readers would perceive the expertise of human-written news as higher than for automated news, proposing the following hypothesis.

H3b: In terms of Chinese readers' actual perceptions, the expertise of human-written news is higher than the expertise of automated news.

In terms of credibility, previous studies have produced mixed results about whether human-written or automated news was rated as more credible. Some studies revealed that stories written by humans were judged as more credible than automated news because writing news is still perceived as a human job (Waddell, 2018), whereas others found automated news was rated slightly more credible when participants read both types of news (Clerwall, 2014; Haim & Graefe, 2017). Overall, even among studies that showed automated news are more credible, the differences between the perceived credibility of the two types of news tend to be small (Graefe et al., 2018; Van der Kaa & Kraemer, 2014). Hereby, this study predicted that the credibility of Chinese automated news would be perceived as inferior to human-written news.

H3c: In terms of Chinese readers' actual perceptions, the credibility of human-written news is higher than the credibility of automated news.

Previous studies suggested different study designs might yield different results. Therefore, this study conducted both within-subject and between-subject design studies to test whether the mode of the experiment will affect results.

Study 1: Method

A 2 (article source: human-written or automated news) x 4 (topic: civic news, sports news, financial reports, conference news) mixed-design experiment was conducted in February of 2019. The article source was a between-subject variable, and the topic was a within-subject variable. Chinese participants ($N = 125$) were randomly assigned to read either human-written ($n = 62$) or automated news ($n = 63$), and each participant read one story on each of the four topics without being informed of the actual source (human or algorithms). The order of the stories was randomized. Differences between topics were not predicted, but four topics were used to test the robustness of the effects. The experiment was embedded in an online survey on Qualtrics. Participants spent an average of 17.23 minutes to complete the experiment.

Participants

A convenience sample of 173 Chinese-speaking participants was recruited through invitations on Chinese online social media platforms including Wechat, Weibo, and Zhihu. After excluding subjects younger than 18 ($n = 1$), repeated IP addresses ($n = 6$), and incomplete answers ($n = 41$), the data set included 125 subjects. Each participant was paid \$0.25. The average age of participants was 23.94 years old ($SD = 3.02$, $Median = 24$). More than half (57.9%) of the participants were female, and 42.1% were male. Participants received 16.73 years of education on average ($SD = 1.92$, $Median = 16$). More than one-third (33.3%) of the participants self-reported that they had read Chinese automated news before the experiment.

Stimuli

Four sets of news articles were selected as the stimuli of the experiment. Each set has two articles of similar lengths and themes, one written by human reporters and the other by algorithms. The researcher first identified 64 pieces of Chinese automated news stories published in 2018 and matched them with human-written news published in the same year that had the most similar topics and length. Finally, four

automated articles with four relatively comparable human-written stories were selected as the stimuli. The final stimuli consisted of four most popular automated news genres in China: sports news, civic news, finance news, and conference news. The sports set included NBA and Inter Milan. The civic news set included two pieces of news covering the Chinese Spring Festival travel rush. The finance news set consisted of two financial reports about local companies. The set of conference news contained news coverage of the highlights of a conference or a talk. It is worth noting that four machine-written stimuli were all fully auto-generated by algorithms. These automated articles explicitly indicated that the author of the news was an algorithm. For instance, the byline of one article reads: "this content is automatically generated by an AI named Xiaonan developed by *Southern Metropolis Daily Media Lab*."

To test whether there was a difference between the characteristics of selected human-written stimuli and automated stimuli, this study used nonparametric Mann-Whitney U test (M-WU) to compare words counts of two types of news, because the word counts data were not normally distributed. Results show no significant difference in words counts ($U = 6.50, p = .66$). Therefore, the stimuli were comparable and would not affect the validity of the experiment. The stimuli were randomly presented to participants to avoid the ordering effect.

Table 1. Feature Analysis of Stimuli in Study 1 and Study 2.

No.	Genre	Source	Published by	Words
1	Civic news	Human	Sanxiang City News	262
2	Civic news	Algorithms	Southern Metropolis Daily Media Lab AI	262
3	Sports news	Human	iFeng.com	894
4	Sports news	Algorithms	Toutiao.com Xiaomingbot	581
5	Finance news	Human	iFeng.com	402
6	Finance news	Algorithms	Southern Metropolis Daily Media Lab AI	334
7	Conference news	Human	World Wide Web	670
8	Conference news	Algorithm	Tencent AI Dreamwriter	874

Measurements

To operationalize prior expectations and actual perceptions, nine descriptors were adapted from Haim and Graefe's (2017) and Clerwall's (2014) descriptors because these have been previously tested in several automated news studies (e.g., Wölker, & Powell, 2018) and translated into Chinese language. By taking the characteristics of Chinese language into account, the descriptors this study chose were coherent, clear, pleasant to read, well written, descriptive, informative, objective, accurate, and trustworthy. Participants rated each descriptor on a 7-point scale with a larger number, meaning the descriptor was more accurate. The measure and scale of prior expectations and actual perceptions were purposely designed to be the same to test whether they were confirmed or not. Prior expectations were measured before participants read an article, and actual perceptions were measured after the experimental manipulation was administered. Descriptors were randomized to reduce the priming effect. To test whether these descriptors work well in the Chinese language, a pretest ($N = 60$) was conducted to test the suitability of the descriptors. Principal component analysis (PCA) with Varimax with a Kaiser normalization rotation method showed that nine descriptors were positively correlated and could be reduced into three main components. The Kaiser-

Meyer-Olkin value (*KMO*) value is .84, and the Bartlett's test was significant ($p < .001$), which proved the descriptors can be used in this study.

Another PCA was used to process the actual data from Study 1. Nine descriptors were highly correlated and could also be reduced to three measurements, namely readability, expertise, and credibility. The *KMO* value is .90, and the Bartlett's test was significant ($p < .001$). Each factor loading was greater than 0.59, and the cumulative variance contribution rate was 80.66%. In summary, the operational definition of readability (*Cronbach's a* = .87, $M = 5.08$, $SD = 3.61$) consisted of three items: coherent, clear, and pleasant to read. The operational definition of expertise (*Cronbach's a* = .83, $M = 4.81$, $SD = 3.49$) included three items: well written, descriptive, and informative. The operational definition of credibility (*Cronbach's a* = .90, $M = 5.30$, $SD = 3.29$) consisted of three items: objective, accurate, and trustworthy.

Study 1: Results

An Analysis of Variance (ANOVA) were used to test H1. H1 predicted that readers' prior expectations of human-written news were higher than automated news, which was partially supported. In terms of readability, participants' prior expectations of human-written news ($M = 5.75$, $SD = .97$) were significantly higher than automated news ($M = 4.80$, $SD = 1.10$), $F(1, 249) = 51.52$, $p < .001$. In terms of expertise, expectations of human-written news ($M = 5.50$, $SD = 1.14$) were significantly higher than automated news ($M = 4.73$, $SD = 1.08$), $F(1, 249) = 37.41$, $p < .001$. In terms of credibility, expectations of human-written news ($M = 4.92$, $SD = 1.15$) were lower than automated news ($M = 5.12$, $SD = 1.18$), but the difference was not statistically significant, $F(1, 249) = 1.88$, $p = .17$. Therefore, H1 was partially supported.

A one-way repeated measures ANOVA was used to test H2. H2a predicted that for human-written news, readers' actual perceptions were lower than their prior expectations, which was partially supported. In terms of readability, readers' perceptions of human-written news ($M = 5.07$, $SD = 1.16$) were significantly lower than their expectations ($M = 5.75$, $SD = .97$), $F(1, 246) = 31.31$, $p < .001$. In terms of expertise, readers' perceptions of human-written news ($M = 4.88$, $SD = 1.41$) were significantly lower than their expectations ($M = 5.50$, $SD = .90$), $F(1, 246) = 31.44$, $p < .001$. In terms of credibility, however, readers' perceptions ($M = 5.23$, $SD = 1.56$) were significantly higher than their expectations ($M = 4.92$, $SD = 1.15$), $F(1, 246) = 6.21$, $p = .01$.

H2b predicted that for automated news, readers' actual perceptions were higher than their prior expectations, which was partially supported. In terms of readability, for automated news, readers' perceptions ($M = 5.03$, $SD = 1.25$) were higher than their expectations ($M = 4.81$, $SD = 1.10$), but not significant, $F(1, 248) = 2.85$, $p = .09$. In terms of expertise, readers' perceptions of automated news ($M = 4.73$, $SD = 1.18$) were the same as their expectations ($M = 4.73$, $SD = 1.08$), $F(1, 248) = .01$, $p = .92$. In terms of credibility, readers' perceptions ($M = 5.36$, $SD = 1.04$) were significantly higher than their expectations ($M = 5.13$, $SD = 1.18$), $F(1, 248) = 3.74$, $p = .05$.

H3 predicted that in terms of Chinese readers' actual perceptions, the (a) readability, (b) expertise, and (c) credibility of human-written news would be higher than the (a) readability, (b) expertise, and (c) credibility of automated news. Multivariate analysis of variance (MANOVA) was used to test H3. The analysis

showed no significant main effect for the source of news articles (human or algorithms) on the perceived readability, $F(4, 120) = 1.91, p = .11, \text{Wilk's } \Lambda = .94, \text{partial } \eta^2 = .06$. No significant main effect was observed on the perceived news credibility, $F(4, 120) = 2.00, p = .10, \text{Wilk's } \Lambda = .94, \text{partial } \eta^2 = .06$. Therefore, H3a and H3c were not supported. However, the analysis revealed a significant main effect for the source of news articles on the perceived news expertise, $F(4, 120) = 5.18, p < .001, \text{Wilk's } \Lambda = .85, \text{partial } \eta^2 = .15$. Specifically, in Story Set 2 (civic news) and Story Set 3 (finance news), the perceived expertise of automated news was significantly lower than that of human-written news, Story Set 2: $F(1, 123) = 4.02, p = .047, \text{partial } \eta^2 = .03$, Story Set 3: $F(1, 123) = 8.49, p = .004, \text{partial } \eta^2 = .07$. In Story Set 4 (conference news), the perceived expertise of automated news was significantly higher than the expertise of human-written news, $F(1, 123) = 4.82, p = .03, \text{partial } \eta^2 = .04$. No significant difference was observed in Story Set 1 (sports news), $F(1, 123) = .21, p = .65, \text{partial } \eta^2 = .002$. Therefore, H3b was partially supported. In sum, no significant differences of readers' actual perception were observed between human-written and automated news in terms of readability and credibility. There existed a significant difference of readers' actual perceptions between human-written and automated news in terms of expertise; however, the direction of the difference depended on the topic of the story, as shown in Table 2.

Table 2. The Mean (SE) Difference between Actual Perceptions of Automated and Human-written News in Study 1 (7-point Scale).

	Readability		Expertise		Credibility	
	Human	Automation	Human	Automation	Human	Automation
Sports	4.89 (.15)	5.18 (.15)	4.94 (.15)	4.84 (.15)	5.42 (.14)	5.70 (.14)
Civic	5.14 (.15)	5.06 (.15)	4.76* (.14)	4.37* (.14)	5.30 (.13)	5.29 (.13)
Finance	5.03 (.15)	4.74 (.15)	5.17** (.15)	4.57** (.14)	5.33 (.14)	5.25 (.13)
Conference	5.23 (.15)	5.13 (.15)	4.67* (.15)	5.13* (.15)	4.88 (.14)	5.20 (.14)

* $p < .05$. ** $p < .01$. *** $p < .001$.

Study 1 examined whether readers' prior expectations and actual perceptions of automated journalism differed from human-written news when they saw only one type of article (human-written or automated news). To sum up, as shown in Figure 1, after reading the article, for human-written news, in terms of readability and expertise, readers' perceptions were significantly lower than their expectations. For automated news, readers' perceptions were higher than their expectations, but only significantly higher in terms of credibility. In terms of expertise, a significant difference was observed between actual perceptions of human-written news and automated news.

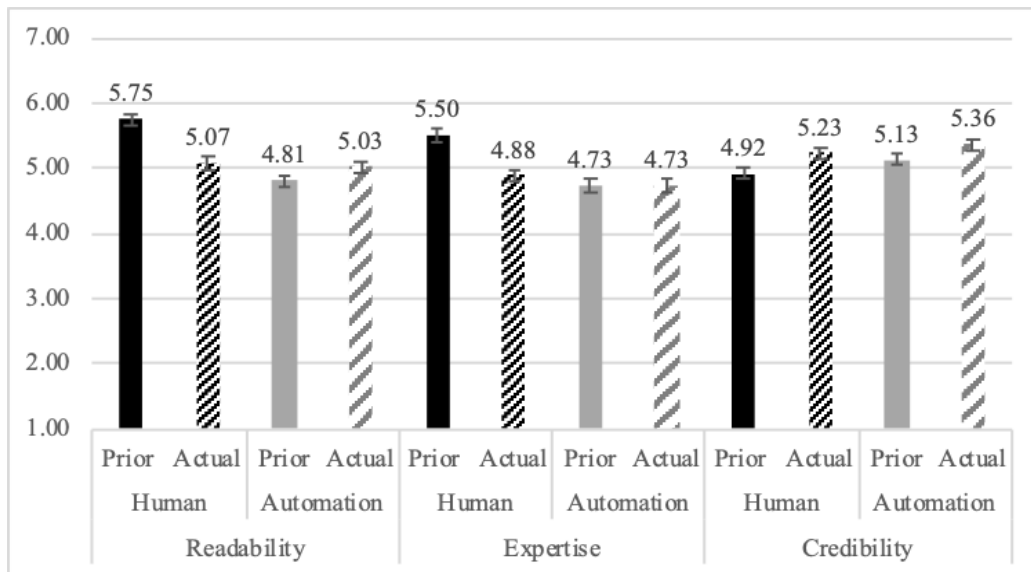


Figure 1. Readers' prior expectations and actual perceptions in study 1 (7-point scale).

Study 2: Method

A 2 (article source: human-written or automated news) x 4 (topic: civic news, sports news, financial reports, conference news) within-subjects design experiment was conducted in April of 2018. Chinese participants ($N = 308$) were asked to read eight articles from four genres without being informed of the actual authorship. Study 2 used the same stimuli as Study 1. The order of the stories was randomized. The experiment was embedded in a Chinese online survey platform named wjx.cn. Participants spent an average of 11.27 minutes to complete the experiment.

Participants

A convenience sample of 313 Chinese-speaking participants was recruited through the same social media platforms used in Study 1. After excluding subjects younger than 18 ($n = 5$), the data set included 308 subjects. The average age of participants was 23.66 years old ($SD = 3.99$, $Median = 23$). More than half (66.9%) of participants were female, and 33.1% were male. Before the experiment, 38.6% of participants had previously read Chinese automated journalism. Participants were entered in a drawing for gift cards provided by the survey platform. More than half (50.9%) of participants had master's degrees and above, and 44.5% participants had received either bachelor's degrees or associate degrees, and 4.6% participants were high school graduates. Before the experiment, 38.3% of participants had read Chinese automated journalism.

Measurements

Participants in Study 2 were required to answer their prior expectations and actual perceptions through a 5-point scale. The measurements were the same as were used in Study 1. Many responses

indicated that a descriptor was more accurate. As in study 1, a PCA indicated the same three factors. These three factors were averaged into indices. Readability (*Cronbach's a* = .86, *M* = 3.63, *SD* = 2.24) consisted of three items: coherent, clear, and pleasant to read. Expertise (*Cronbach's a* = .80, *M* = 3.48, *SD* = 2.21) consisted of three items: well-written, descriptive, and informative. Credibility (*Cronbach's a* = .87, *M* = 3.61, *SD* = 2.14) consisted of three items: objective, accurate, and trustworthy. The *KMO* value was .89, and the Bartlett's test was significant ($p < .001$). Each factor loading was greater than .55, and the cumulative variance contribution rate was 78.07%.

Manipulation Check

To ensure the success of the experimental design, after reading each news article, a manipulation check question was asked to determine whether participants could discern whether the article they had just read was written by a human or an algorithm. A paired sample t-test showed that the manipulation check resulted in a significant difference between the two conditions (human or algorithms), $t(307) = 3.71, p < .001$, which suggested that the manipulation was a success.

Study 2: Results

An Analyses of variance (ANOVA) was used to test H1. H1 predicted that readers' expectations of human-written news were higher than automated news. In terms of readability, participants' expectations of human-written news ($M = 4.02, SD = .67$) were significantly higher than automated news ($M = 3.32, SD = .66$), $F(1, 615) = 164.18, p < .001$. In terms of expertise, participants' expectations of human-written news ($M = 3.85, SD = .65$) were significantly higher than automated news ($M = 3.19, SD = .70$), $F(1, 615) = 147.06, p < .001$. In terms of credibility, participants' expectations of human-written news ($M = 3.46, SD = .72$) were lower than automated news ($M = 3.49, SD = .80$), but the difference was not statistically significant, $F(1, 615) = .14, p = .71$. Therefore, H1 was partially supported, and the results yielded the same pattern as in Study 1.

A one-way repeated measures ANOVA was used to test H2. H2a predicted that for human-written news, readers' actual perceptions were lower than their prior expectations. In terms of readability, for human-written news, readers' perceptions ($M = 3.82, SD = .66$) were significantly lower than their expectations ($M = 4.02, SD = .67$), $F(1, 614) = 15.94, p < .001$. In terms of expertise, readers' perceptions ($M = 3.64, SD = .65$) were significantly lower than their expectations ($M = 3.85, SD = .65$), $F(1, 614) = 20.80, p < .001$. In terms of credibility, however, readers' perceptions ($M = 3.74, SD = .65$) were significantly higher than their expectations ($M = 3.46, SD = .72$), $F(1, 614) = 31.02, p < .001$. Therefore, H2a was partially supported. These results were also the same as Study 1.

H2b predicted that for automated news, readers' actual perceptions were higher than their prior expectations. In terms of readability, for automated news, readers' actual perceptions ($M = 3.40, SD = .77$) were higher than their prior expectations ($M = 3.32, SD = .66$), but not significant, $F(1, 614) = 2.60, p = .11$. In terms of expertise, readers' perceptions ($M = 3.31, SD = .77$) were significantly higher than their expectations ($M = 3.19, SD = .70$), $F(1, 614) = 5.63, p = .01$. In terms of credibility, readers' perceptions

($M = 3.56$, $SD = .74$) were higher than their expectations but not statistically significant ($M = 3.49$, $SD = .80$), $F(1, 614) = 1.87$, $p = .17$. Therefore, H2b was partially supported.

Multivariate analysis of variance (MANOVA) was used to test H3. The analysis revealed a significant main effect for the source of news articles (human or algorithm) on the perceived news readability, $F(4, 611) = 35.74$, $p < .001$, *Wilk's* $\Lambda = .81$, partial $\eta^2 = .19$. Specifically, in each set, the readability of automated news was significantly inferior to the readability of human-written news, Story Set 1: $F(1, 614) = 27.37$, $p < .001$, partial $\eta^2 = .04$; Story Set 2: $F(1, 614) = 83.53$, $p < .001$, partial $\eta^2 = .12$; Story Set 3: $F(1, 614) = 108.24$, $p < .001$, partial $\eta^2 = .15$; Story Set 4: $F(1, 614) = 22.24$, $p < .001$, partial $\eta^2 = .04$. H3a was fully supported. The analysis also yielded a significant main effect for the source of news articles on the perceived news expertise, $F(4, 611) = 50.63$, $p < .001$, *Wilk's* $\Lambda = .75$, partial $\eta^2 = .25$, but not every story set had a significant difference between human-written news and automated news. In Story Set 1 (sports news), Story Set 2 (civic news), and Story Set 3 (finance news), the expertise of automated news was significantly inferior to the expertise of human-written news, Story Set 1: $F(1, 614) = 31.27$, $p < .001$, partial $\eta^2 = .05$; Story Set 2: $F(1, 614) = 124.40$, $p < .001$, partial $\eta^2 = .17$; Story Set 3: $F(1, 614) = 64.61$, $p < .001$, partial $\eta^2 = .10$. For Story Set 4 (conference news), however, the expertise of automated news was higher than the expertise of human-written news, but not significantly, $F(1, 614) = .05$, $p = .83$, $\eta^2 = .00$. Therefore, H3b was partially supported. For H3c, a significant main effect for the source of news articles on the perceived news credibility was observed, $F(4, 611) = 65.67$, $p < .001$, *Wilk's* $\Lambda = .70$, partial $\eta^2 = .30$, but not every story set revealed the same direction as hypothesized, as shown in Table 3. Only in Story Set 2 (civic news), the credibility of automated news was significantly lower than the credibility of human-written news, $F(1, 614) = 192.92$, $p < .001$, partial $\eta^2 = .24$, as hypothesized. Therefore, H3c was not supported.

Table 3. The Mean (SE) Difference between Actual Perceptions of Automated and Human-written News in Study 2 (5-point Scale).

	Readability		Expertise		Credibility	
	Human	Automation	Human	Automation	Human	Automation
Sports	3.91*** (.04)	3.62*** (.04)	3.82*** (.04)	3.54*** (.04)	3.70 (.04)	3.74 (.04)
Civic	3.76*** (.04)	3.22*** (.04)	3.39*** (.04)	2.75*** (.04)	3.97*** (.04)	3.17*** (.04)
Finance	3.89*** (.04)	3.31*** (.04)	3.82*** (.04)	3.41*** (.04)	3.72 (.04)	3.69 (.04)
Conference	3.74*** (.04)	3.47*** (.04)	3.52 (.04)	3.53 (.04)	3.57 (.04)	3.65 (.04)

* $p < .05$. ** $p < .01$. *** $p < .001$.

As shown in Figure 2, results in Study 2 showed that readers' actual perceptions of human-written news were significantly lower than their expectations in readability and expertise, but significantly higher in credibility. Readers' actual perceptions of automated news significantly exceeded their expectations in expertise. For both civic and financial news, actual perceptions of human-written news were superior to automated news in terms of readability, expertise and credibility.

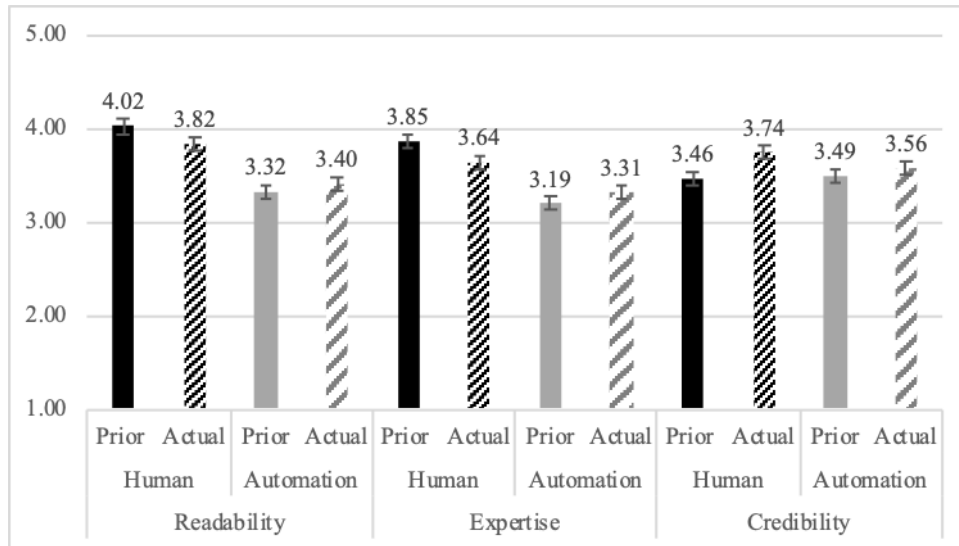


Figure 2. Readers' prior expectations and actual perceptions in study 2 (5-point scale).

Discussion

The current study broadened the geographical range of research on automated journalism through examining readers' prior expectations and actual perceptions of automated news written in Chinese. News is a cultural product restricted not only by language but also by cultural boundaries (Chyi et al., 2019). Chinese automated journalism is a cutting-edge technology born in this market. Previous studies suggested cultural background matters when Chinese users and American users perceived the content purportedly written by or humans or algorithms (Zheng et al., 2018). However, earlier studies have not fully examined the perceived quality of automated news written in different languages. For instance, Zheng et al. (2018) used the same content and manipulated the byline in two conditions (human or algorithm). Their study was not capable of mirroring the true quality of Chinese automated news because stimuli used by them were originally generated by English algorithm providers and then were translated into Chinese. Similarly, Clerwall's (2014) study asked Swedish undergraduate students to read stories written in English with the topic of American football, which was not representative of native English news consumers. Adding to these literatures, the present study examined how Chinese users perceive content generated by local algorithm providers.

Results showed that when participants read both human-written and automated news, actual perceptions of human-written news were significantly higher than automated news in terms of readability and expertise. When participants read either human-written news or automated news, significant difference existed only in expertise. No significant difference in credibility was observed in either study. These findings were consistent with several previous studies. For instance, participants favored human-written news in terms of readability (Clerwall, 2014; Haim & Graefe, 2017), and perceived the credibility and trustworthiness of human-written and automated news equally (Graefe et al., 2018; Van der Kaa & Kraemer, 2014; Zheng et al., 2018).

Another major finding is that for human-written news, the perceived readability and expertise were significantly lower than expectations in both studies whereas the perceived credibility was significantly higher than expected in Study 1. In other words, human-written news was always overrated in the prior expectations of readability and expertise but was underrated in terms of credibility. For automated news, readers' actual perceptions were higher overall than their prior expectations but only significant in the credibility measure in Study 1 and expertise measure in Study 2. This current study demonstrated human-written news did not meet people's expectations, which is consistent with Haim and Graefe's (2017) study. On the contrary, the finding of this study confirmed that readers' perceptions of automated news were higher than their expectations, at least in certain measures; whereas, Haim and Graefe (2017) found that automated news did not meet up with people's expectations. The difference could possibly be explained by the distinct cultural contexts. Haim and Graefe (2017) conducted this study in Germany. Past work found that German participants perceived sociable robots less likable, trustworthy, and satisfactory than Chinese participants (Li, Patrick Rau, & Li, 2010). Aside from this, past media use and experience could affect and shape audiences' perceptions of automation technology (Sundar et al., 2016). The present work reported the percentages of participants who had read automated news before the experiment (33.3% in Study 1 and 38.3% in Study 2). Haim and Graefe's (2017) work did not report that information so no comparison can be made in terms of past experience. Future studies can reference past media use as a supplemental measure to examine ECT or other theories such as the expectancy violation theory (e.g., Burgoon, 2015).

Another contribution of this study is that comparisons of actual perceptions were made across different news genres. When participants read either human-written news or automated news, a significant difference only existed in the perceived expertise. The direction of the difference, however, depended on the topic of the story. The expertise of human-written news was significantly higher in the realm of civic news and financial news and was lower in conference stories. No significant difference was observed in the area of sports news. When participants read both human-written and automated news, actual perceptions of human-written news were significantly higher than automated news in terms of readability (in every topic) and expertise (except for conference stories). In both studies, the expertise of human-written civic news and finance news were rated significantly higher than that of automated news. A possible explanation might be that readers require more professional knowledge when reading financial and civic-related news. There are usually many statistical numbers in the financial or civic news. Readers might expect the author to establish causality or explain the underlying reasons behind those numbers, rather than just citing or extracting numbers from data sets. Human authors usually explicate numbers, whereas algorithms tend to pile on numbers without explaining because of the limitations of their working principles. Sports articles were rated similarly in terms of readability and expertise when participants read only one type of news. However, when participants read both types of stories, the perceived readability and expertise of human-written sports news were both rated significantly higher than automated news. Although the results might be affected by study design, it is noteworthy that algorithm providers in China have achieved a breakthrough to make automated sports news more diverse and livelier. In China, automated journalism in sports is no longer restricted to prewritten templates. Instead, more advanced natural language generation technologies or facial recognition technologies are applied to sports coverage. For instance, the *Xinhua News Agency* used algorithms to capture the exaggerated facial expressions of football coaches, and then auto-generated sports news videos during the 2018 FIFA World Cup (Jia, 2018). Therefore, it is not surprising if the differences between human-written news and automated sports news are not as obvious as in civic and finance news.

Cultural Context of Chinese Automated Journalism

Chinese automated journalism started relatively late (in 2015) but has developed rapidly because of the enormous market. Three years after Tencent published its first piece of Chinese automated news, the algorithm of Tencent automated journalism iterates every three months on average (K. Liu, personal communication, April 24, 2018). According to K. Liu, Project Director of Tencent Dreamwriter, Tencent autogenerated around 2,000 financial news stories and at least 500 pieces of sports news per day in 2018. Alongside Internet companies, state-run *Xinhua News Agency* also began to automate sports news, civic news, and even political news content such as coverage of the National People's Congress (Xinhuanet, 2018).

Both Study 1 and Study 2 found that Chinese readers had higher prior expectations for human-written news than automated news in terms of readability and expertise, but the expectations for credibility of the two types of news were similar. Moreover, the actual perceived credibility of two types of stories were rated similar after exposure to articles. With that said, the credibility of automated and human-written news was again proven to be indiscernible. Although the similarity in credibility was consistent with previous work (Clerwall, 2014; Graefe et al., 2018; van der Kaa & Kraemer, 2014; Zheng et al., 2018), it can be explained by the uniqueness of Chinese media and cultural sensibility. In both studies, readers' expectations for credibility of human-written news was the lowest among all three measurements. One possible explanation is that, compared with Western countries' market-oriented media system, the Chinese media system is relatively centralized and politically oriented (Zhang, Zhou, & Shen, 2014). Chinese people perceive state-run media outlets to be more credible (e.g., China Central Television) than other media (Zhang et al., 2014). This study, however, only used the vague phrase "human journalists" without revealing the name of specific media outlets, which might have weakened the perceived credibility of human-written news. An alternative explanation is that Chinese participants perceived social robots as more trustworthy and had higher engagement with the AI technology compared with Western participants (Li et al., 2010). It is possible that Chinese-speaking participants had similar expectations for automated and human-written news because they were more open to automation technology and had relatively higher trust in the algorithm. It is good news for the Chinese automated news market because higher trust in algorithms can be an impetus for the effective adoption of automation technology (Lee & See, 2004; Lee, 2018).

Limitations and Future Research

Despite these contributions to the automated journalism studies, this work had certain limitations. First, this study only tested and confirmed part of the ECT model's rationale. Because of the design of this study, participants were not informed of the actual source of news they read (i.e., human or algorithm). It would be fallacious to ask them to measure satisfaction after reading automated journalism and human-written news if they did not know the actual sources of a given article. Future studies can test the full model by measuring postadoption satisfaction and continuous intention after participants read the news and are informed of the actual authorship. According to the results of Study 1, readers' perceptions of automated news were significantly higher than expectations in terms of credibility. Would this lead to readers' satisfaction in consuming automated news? Future studies may measure participants' intentions to read the automated news in the future.

Second, two experiments in this study used two scales and different survey platforms. Even though these differences did not affect the internal validity of this study, it would be better for future studies to use comparable study designs to reduce unnecessary confounding factors. Third, participants in both studies were relatively young. Although young readers are likelier to encounter automated news than the older generations, future research should be more rigorous in sampling. Lastly, this study adapted Haim and Graefe's (2017) and Clerwall's (2014) descriptors and translated them into Chinese. Although this study tested the suitability of these descriptors in a Chinese context by conducting a pretest analysis, the generalizability of these descriptors and their three main components remains uncertain. Overall, even though research on automated journalism are still in nascent stages, this area can be a fertile field for future research because of the sheer number of individuals exposed to automated news products.

References

- Bhattacharjee, A. (2001). Understanding information systems continuance: An expectation-confirmation model. *MIS Quarterly*, 25(3), 351–370. <https://doi.org/10.2307/3250921>
- Burgoon, K. J. (2015). Expectancy violations theory. In C. R. Berger & M. E. Roloff (Eds.), *The international encyclopedia on interpersonal communication* (pp. 1–9). New York, NY: Wiley Blackwell.
- Carlson, M. (2015). The robotic reporter. *Digital Journalism*, 3(3), 416–431. <https://doi.org/10.1080/21670811.2014.976412>
- Carlson, M. (2017). *Journalistic authority: Legitimizing news in the digital era*. New York, NY: Columbia University Press.
- Chung, C., Nam, Y., & Stefanone, M. (2012). Exploring online news credibility: The relative influence of traditional and technological factors. *Journal of Computer-Mediated Communication*, 17(2), 171–186. doi:10.1111/j.1083-6101.2011.01565.x
- Chung, M. (2017). Not just numbers: The role of social media metrics in online news evaluations. *Computers in Human Behavior*, 75(C), 949–957. doi:10.1016/j.chb.2017.06.022
- Chyi, H. I., Lee, M. A., & Tennant, I. J. (2019). News across the great wall: Asian news organisations' Web strategies for the China market. *Journal of Media Business Studies*, 16(4), 307–325. <https://doi.org/10.1080/16522354.2019.1682906>
- CIA. (2016). The world factbook—Central Intelligence Agency. Retrieved from <https://www.cia.gov/library/publications/the-world-factbook/rankorder/2153rank.html#in>
- Clerwall, C. (2014). Enter the robot journalist. *Journalism Practice*, 8(5), 519–531. <https://doi.org/10.1080/17512786.2014.883116>

- Dalecki, L., Lasorsa, D. L., & Lewis, S. C. (2009). The news readability problem. *Journalism Practice*, 3(1), 1–12. <https://doi.org/10.1080/17512780802560708>
- Diakopoulos, N. (2015). Algorithmic accountability. *Digital Journalism*, 3(3), 398–415. doi:10.1080/21670811.2014.976411
- Domingos, P. (2015). *The master algorithm: How the quest for the ultimate learning machine will remake our world*. St. Ives, UK: Allen Lane.
- Dörr, K. N. (2016). Mapping the field of algorithmic journalism. *Digital Journalism*, 4(6), 700–722. doi:10.1080/21670811.2015.1096748
- Dubay, W. (2004). *The principles of readability*. Costa Mesa, CA: Impact Information.
- Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford, CA: Stanford University Press.
- Fountas, I. C., & Pinnell, G. S. (1999). *Matching books to readers: Using leveled books in guided reading*. Portsmouth, NH: Heinemann.
- Fu, P., Shang, Y., & Zhang, N. (2019). The evolution of intellectual media in the “media brain”: Everything media and human-computer symbiosis. Retrieved from <http://media.people.com.cn/n1/2019/0409/c426142-31019962.html>
- Graefe, A. (2016). Guide to automated journalism. *Tow Center for Digital Journalism*. New York, NY: Columbia Journalism School. <https://doi.org/10.7916/D80G3XDJ>
- Graefe, A., Haim, M., Haarmann, B., & Brosius, H.-B. (2018). Readers’ perception of computer-generated news: Credibility, expertise, and readability. *Journalism*, 19(5), 595–610. <https://doi.org/10.1177/1464884916641269>
- Hagar, N., & Diakopoulos, N. (2019). Optimizing content with a/b headline testing: Changing newsroom practices. *Media and Communication*, 7(1), 117–127. doi:10.17645/mac.v7i1.1801
- Haim, M., & Graefe, A. (2017). Automated news: Better than expected? *Digital Journalism*, 5(8), 1044–1059. <https://doi.org/10.1080/21670811.2017.1345643>
- Heider, F. (1946). Attitudes and cognitive organization. *Journal of Psychology*, 21(1), 107–112. <https://doi.org/10.1080/00223980.1946.9917275>
- Hovland, C. I., Janis, I. L., & Kelley, H. H. (1953). *Communication and persuasion*. New Haven, CT: Yale University Press.

- Jia, C. (2018). AI generated over 10 thousand FIFA World Cup videos in one week. Only takes 16 seconds to auto-generate one video. Retrieved from <http://www.ciweigongshe.net/index/detail?id=734&type=0>
- Jung, J., Song, H., Kim, Y., Im, H., & Oh, S. (2017). Intrusion of software robots into journalism: The public's and journalists' perceptions of news written by algorithms and human journalists. *Computers in Human Behavior, 71*, 291–298. <https://doi.org/10.1016/j.chb.2017.02.022>
- LeCompte, C. (2015). Automation in the newsroom—How algorithms are helping reporters expand coverage, engage audiences, and respond to breaking news. *Nieman Journalism Lab*. Retrieved from <https://niemanreports.org/articles/automation-in-the-newsroom/>
- Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors, 46*(1), 50–80. https://doi.org/10.1518/hfes.46.1.50_30392
- Lee, M. K. (2018). Understanding perception of algorithmic decisions: Fairness, trust, and emotion in response to algorithmic management. *Big Data & Society, 5*(1), 1–16. <https://doi.org/10.1177/2053951718756684>
- Li, D., Patrick Rau, P. L., & Li, Y. (2010). A cross-cultural study: Effect of robot appearance and task. *International Journal of Social Robotics, 2*, 175–186. <https://doi.org/10.1007/s12369-010-0056-9>
- Lin, C. S., Wu, S., & Tsai, R. J. (2005). Integrating perceived playfulness into expectation-confirmation model for Web portal context. *Information & Management, 42*(5), 683–693. <https://doi.org/10.1016/j.im.2004.04.003>
- Lokot, T., & Diakopoulos, N. (2016). News bots: Automating news and information dissemination on twitter. *Digital Journalism, 4*(6), 682–699. doi:10.1080/21670811.2015.1081822
- Mager, A. (2012). Algorithmic ideology: How capitalist society shapes search engines. *Information, Communication & Society, 15*(5), 769–787. doi:10.1080/1369118X.2012.676056
- Metzger, M. J., Flanagin, A. J., Eyal, K., Lemus, D. R., & McCann, R. (2003). Credibility in the 21st century: Integrating perspectives on source, message, and media credibility in the contemporary media environment. In P. Kalbfleisch (Ed.), *Communication yearbook 27* (pp. 293–335). Mahwah, NJ: Lawrence Erlbaum. <https://doi.org/10.1080/23808985.2003.11679029>
- Meyer, P. (2004). *The vanishing newspaper: Saving journalism in the Information Age*, Columbia: University of Missouri Press.

- Napoli, P. M. (2014). Automated media: An institutional theory perspective on algorithmic media production and consumption. *Communication Theory, 24*(3), 340–360. <https://doi.org/10.1111/comt.12039>
- Newcomb, T. M. (1959). Individual systems of orientation. In S. Koch (Ed.), *Psychology: A study of a science 3: Formulations of the person and the social context* (pp. 384–422). New York, NY: McGraw-Hill.
- Noack, R., & Gamio, L. (2015). The world's languages, in 7 maps and charts. Retrieved from https://www.washingtonpost.com/news/worldviews/wp/2015/04/23/the-worlds-languages-in-7-maps-and-charts/?noredirect=on&utm_term=.ff6fc3172b3f
- Oliver, R. L. (1980). A cognitive model of the antecedents and consequences of satisfaction decisions. *Journal of Marketing Research, 17*(4), 460–469. <https://doi.org/10.2307/3150499>
- Oliver, R. L. (1993). Cognitive, affective, and attribute bases of the satisfaction response. *Journal of Consumer Research, 20*(3), 418–430. <https://doi.org/10.1086/209358>
- Oliver, R. L. (2015). *Satisfaction. A behavioral perspective on the consumer* (2nd ed.). London, UK: Routledge.
- Oliver, R. L., & DeSarbo, S. W. (1988). Response determinants in satisfaction judgments. *Journal of Consumer Research, 14*(4), 495–507. <https://doi.org/10.1086/209131>
- Slater, M. D., & Rouner, D. (1996). How message evaluation and source attributes may influence credibility assessment and belief change. *Journalism & Mass Communication Quarterly, 73*(4), 974–991. <https://doi.org/10.1177/107769909607300415>
- Sun, W., Chen, Y., Wan, X., & Liu, M. (2019). Parsing Chinese sentences with grammatical relations. *Computational Linguistics, 45*(1), 95–136. doi:10.1162/coli_a_00343
- Sundar, S. S. (1999). Exploring receivers' criteria for perception of print and online news. *Journalism & Mass Communication Quarterly, 76*(2), 373–386. doi:10.1177/107769909907600213
- Sundar, S. S., Waddell, T. F., & Jung, E. (2016, March). The Hollywood robot syndrome: media effects on older adults' robot attitudes and adoption intentions. In *Proceedings of HRI 2016: ACM/IEEE International Conference on Human-Robot Interaction*, 343–350. Christchurch, New Zealand. doi:10.1109/HRI.2016.7451771
- Thurman, N. (2011). Making "the daily me": Technology, economics and habit in the mainstream assimilation of personalized news. *Journalism: Theory, Practice & Criticism, 12*(4), 395–415. <https://doi.org/10.1177/1464884910388228>

- Thurman, N., Dörr, K., & Kunert, J. (2017). When reporters get hands-on with robo-editing: Professionals consider automated journalism's capabilities and consequences. *Digital Journalism*, 5(10), 1240–1259. <https://doi.org/10.1080/21670811.2017.1289819>
- Van der Kaa, H., & Kraemer, E. (2014, October). Journalist versus news consumer: The perceived credibility of machine written news. In *Proceedings of the Computation + Journalism Conference*. New York, NY.
- Waddell, T. F. (2018). A robot wrote this? How perceived machine authorship affects news credibility. *Digital journalism*, 6(2), 236–255. <https://doi.org/10.1080/21670811.2017.1384319>
- Waddell, T. F. (2019). Can an algorithm reduce the perceived bias of news? Testing the effect of machine attribution on news readers' evaluations of bias, anthropomorphism, and credibility. *Journalism & Mass Communication Quarterly*, 96(1), 82–100. <https://doi.org/10.1177/1077699018815891>
- Wölker, A., & Powell, T. E. (2018). Algorithms in the newsroom? News readers' perceived credibility and selection of automated journalism. *Journalism*, 1–18. <https://doi.org/10.1177/1464884918757072>
- Xinhuanet. (2018). Human vs. AI: How Xinhua's media brain covers the two sessions. Retrieved from http://www.xinhuanet.com/english/2018-03/22/c_137056423.htm
- Yin, T. (2017). The arrival of robot journalism age: Toutiao.com, Tencent and Southern Weekend adopt robot journalists, media workers face layoffs. Retrieved from <http://www.iheima.com/zixun/2017/0719/164208.shtml>
- Zhang, H., Zhou, S., & Shen, B. (2014). Public trust: A comprehensive investigation on perceived media credibility in China, *Asian Journal of Communication*, 24(2), 158–172. doi:10.1080/01292986.2013.856452
- Zheng, Y., Zhong, B., & Yang, F. (2018). When algorithms meet journalism: The user perception to automated news in a cross-cultural context. *Computers in Human Behavior*, 86, 266–275. <https://doi.org/10.1016/j.chb.2018.04.046>

