

Believing Journalists, AI, or Fake News: The Role of Trust in Media

Castulus Kolo
Macromedia University
of Applied Sciences
c.kolo@macromedia.de

Joschka Mütterlein
Macromedia University
of Applied Sciences
j.muetterlein@macromedia.de

Sarah Anna Schmid
Macromedia University
of Applied Sciences
s.schmid@macromedia.de

Abstract

An increasing amount of news is generated automatically by artificial intelligence (AI). While the technology has advantages for content production, e.g., regarding efficiency in aggregating information, it is also viewed critically due to little transparency in obtaining results and possible biases. As news media are dependent on trust and credibility, introducing AI to facilitate mass communication with consumers seems to be a risky endeavor. We expand research on consumer perception of AI-based news by comparing machine-written and human-written texts to fake news and by examining the role of trust that consumers exhibit when evaluating news. Through an experiment with 263 participants, we find that consumers judge AI-based texts similar to true journalistic content when it comes to credibility, but similar to fake news regarding readability. Furthermore, our results indicate that consumers with low trust in media are less averse to AI-based texts than consumers with high trust in media.

1. Introduction

Artificial Intelligence (AI) is developing rapidly and gets applied in a many different areas [1]. The changes that AI will bring to our everyday life are already noticeable. This can also be seen in the media industry, which is increasingly shaped by digitalization and AI technologies. The latter are used for, e.g., automated text generation, translation of texts, and information research. They can greatly facilitate the editorial work and offer possibilities to create new digital services [2]. Furthermore, AI is a potential solution to support and simplify journalistic value creation processes at various points [3]. Especially in content production, AI helps in data aggregation and in identifying relevant information. This can significantly change how mass communication takes place and is perceived.

One of the AI technologies driving this change is Natural Language Processing (NLP). NLP is capable of processing and interpreting human language [4]. A subfield of NLP, which specifically deals with the

creation of text content, is called Natural Language Generation (NLG). The goal of NLG is to produce natural language text based on non-linguistic representation of information [5, 6]. In journalism, NLG often comes in the form of data-to-text generation. Recurring information or data on current developments are imported into existing text modules, for example in weather or soccer reports.

Considering the fast developments of AI in the field of NLG technology, media companies will be able to generate much more complex articles for journalism automatically in the future. Thus, their conversations with consumers will become more dependent on NLG. Whether consumers are willing to take part in such conversations or start to reject such offers of media companies depends on consumers' perspectives on automated communication of news [2]. Researchers have already started to investigate these consumer perspectives, with results indicating that computer-generated texts tend to have the same or a slightly higher credibility, but the readability of AI-generated texts with the current state of technology is still lagging human-written texts [e.g., 2, 7, 8].

In our study, we build on this knowledge and advance it in two ways: First, we investigate consumers' general trust in media as an important personality trait that helps to paint a more nuanced picture of when audiences find AI-generated media content credible. In journalism, trust is an important foundation for consumers to believe the news a media company offers and to keep participating in the (mass-)conversation with this media company. When it comes to AI, trust is even more crucial, because consumers—and often even the developers—of AI-based products cannot judge how an AI comes to its conclusions [9]. Second, we expand the previously often used dichotomy of human-written vs. AI-created texts (two categories sometimes also combined with each other [2]) with fake news as a third category highly relevant in the context of credibility and trust. Integrating fake news in our research, i.e. false or misleading news [10], allows us to place the credibility of AI-generated news on a spectrum ranging from objectively well-written journalistic content [11] to

objectively false content [10] and to judge consumers' perception of texts written by an AI in relation to these two major categories of news content that are shaping or threatening societies worldwide [12].

In the following, to shed light on how consumers with varying levels of trust perceive AI in media, we first expand previous research by comparing consumers' view on true journalistic content, fake news, and an AI-based text. We then examine the influence of media trust on the credibility, readability, and perception of journalistic expertise in texts that are either not labeled or labeled as written by an AI. These two experimental analyses help to answer the following research question:

How does trust in media influence consumers' perception of AI-generated media content?

We collected data for our analyses with a 3x2 between-subjects experiment and 263 participants. In each group, participants were shown one of three possible kinds of texts (original journalistic content, fake news, AI-generated text) that were additionally either labeled as AI-generated or not labeled at all.

Our findings show that consumers find journalistic and AI-generated texts comparable and mostly superior to fake news when it comes to credibility and journalistic expertise, but fake news and AI-generated texts are equally worse than journalistic texts when it comes to readability. In addition, our results indicate that labeling texts as written by an AI decreases their perceived readability and possibly also their perceived credibility and journalistic expertise for consumers with a generally high trust in media, but we could find no such effects or tendencies for consumers with generally low trust in media.

This implicates that companies working in markets sensitive to trust, such as journalism, cannot use AI-generated products for conversations with all their customers and expect the same outcome, but a more targeted approach is needed. Whilst AI-based texts might be perceived not so well by trusting readers, readers with lower levels of trust might not worry at all. Thus, differentiating customers and application areas for AI is necessary for a successful introduction of AI in communication. For research, our results demonstrate that investigating the perception of AI-based news benefits from introducing fake news as third category besides machine-written and human-written content. In the context of trust and credibility, this allows for placing consumer perceptions of AI-based news on a scale and comparing it to the most and least desirable categories media companies want their content to fall into. In addition, we show that consumers' general trust in media is an important variable to consider and at least control for when evaluating perceptions of AI-based news.

2. Theoretical Foundations

2.1 Natural Language Generation in News

According to the most-cited definition of NLG, the term "natural language generation" relates to "the subfield of AI and computational linguistics that is concerned with the construction of computer systems that can produce understandable texts in English or other human languages from some underlying non-linguistic representation of information" [6].

Existing applications of NLG can be divided into two types according to the form of input data. Text-to-text generation applications use existing man-made text modules as input data and automatically produce new texts as output. A more common method for text generation within NLG is data-to-text generation. For data-to-text generation information based on non-linguistic data is used to produce reports. The form of the input data can vary greatly, from simple figures on current developments to information from pictures [13]. In most cases, NLG applications are still based on a template-based approach, which transfers recurring information or data on current developments into existing text modules. Advanced systems already exist, but these still do not resemble human creative writing processes.

Yet, AI, and especially the sub-area of automated content creation using NLG, offers new opportunities for media companies to create competitive advantages in a highly competitive and rapidly changing market environment by increasing efficiency and effectiveness of content production. In the conventional journalistic value-added process, relevant information is aggregated via various channels by the editor. Due to the increasing amount of data, it is becoming more and more difficult for editors to identify and combine relevant information in real time. This results in delays between the occurring event and the moment of reporting. Journalists spend an increasing amount of their working time to check and confirm primary sources instead of writing high-quality content [14] or proceeding with further-reaching investigations. In such circumstances, AI can take over the task of data aggregation and identify potentially interesting information as well as interrelations within it. This can increase the speed as well as the overall quality of journalistic work.

The advantages of NLG can also be seen in text production. Currently, recurring news articles such as weather or sports news are mainly produced automatically, thus relieving journalists of routine tasks [15]. It is to be expected that advantages in NLG technology will soon allow for more complex content to be created.

2.2 AI-generated News, Credibility, and Trust

In automated text generation, the perception of news quality by consumers plays an important role for establishing the relationship between humans and AI. However, the quality assessment of news articles is complex. For example, credibility as one of the most important dimensions of content quality in news [16] depends on many things besides the content, two of which are the trustworthiness of the specific communication source [17] and the medium in general through which content is consumed [18]. These insights are valid and relevant beyond the media industry. As communication is not only increasingly based on AI and NLG in news production, but other industries are more and more following suit, e.g., through chatbots and recommendation systems [e.g., 19], credibility of AI-based text creation and distribution and trust in AI-based texts are highly relevant subjects for research in IS.

To elucidate the current state of research in this regard and the methods commonly used, we briefly summarize three studies on AI-generated texts. One of the earliest studies is by Clerwall [16], who evaluated the perceived quality of news articles using the factors credibility and readability. The 46 participants in the study were presented with one of two articles on the topic of American football without reference to the news source. One of these articles was generated by an AI, the second by a journalist. After being asked about the news quality of the presented text, participants were required to give an estimate of whether the text was written by an AI or a human. Differences in the quality assessment of the AI-generated and journalistic article were small.

To investigate whether the perception of quality changes when news consumers know the source of the article, or at least think they do, van der Kaa and Kraemer [8] investigated the perceived credibility, measured as trustworthiness and journalistic expertise, of computer-generated content. The 232 participants (including journalists and news consumers) in the study were shown one of two computer-generated articles about financial news or the results of a sports event. The articles were manipulated in their byline and were either correctly declared as AI-generated or erroneously declared as written by a journalist. Differences in the perceived quality by news consumers were rather small, but contents with AI declaration were rated slightly better. No differences could be found in the perceived quality for the interviewed journalists. Regarding the articles' content, small differences in the perceived trustworthiness could be shown.

Based on the two studies, Graefe et al. [7] addressed the question of whether news consumers are actually influenced by the declaration of the source and, thus, their expectation of the article's perceived quality. In

addition, they examined how computer-generated news articles and articles written by humans with correct declaration of the source differ in their perceived quality. Results of this study confirmed the results of previous studies. The credibility of AI-generated articles tended to be rated higher than that of human-written texts. Readability was rated higher for human-written texts than for computer-written texts while differences in the perception of credibility were small.

As exemplified with these three studies, research indicates that computer-generated texts tend to have the same or a slightly higher credibility as human-written content, but the readability of AI-generated texts is still lagging human-written texts. Other researchers are supporting these findings or come to similar conclusions. For example, Wölker and Powell [2] identify perceptions of credibility of human-written and automatically created content as well as combinations thereof as equal and mention topic-specific factors as important to consider. Finally, a recent meta-analysis on the topic found that there is no difference in consumers' perception of credibility when comparing automated and human-written news, but the latter performed slightly better in terms of perceived quality and much better regarding readability [20].

Our study builds on these insights and expands them first by incorporating trust. In the context of media, trust means that consumers are willing to rely on information provided by a sender [21]. The concept of credibility is closely related, referring to the believability of a sender [22]. As described above, both concepts are highly relevant for companies or brands respectively in the field of journalism, as they need to appear credible and maintain consumers' trust to ensure that their information will also be consumed in the future [23]. Through incorporating consumers' general trust in media in our study, we can determine whether consumers with higher or lower levels of trust tend to believe AI-generated texts more.

Second, our study expands the mentioned insights by introducing fake news as a third category important in the context of news credibility and consumer trust in news. The term fake news relates to information that appear journalistic but are objectively false or misleading [10] and, thus, rationally not believable. Introducing fake news as a third category allows us to assess to which extent consumers trust AI-generated content along the spectrum ranging from true journalistic news on the one side to objectively false fake news on the other side.

2.3 Hypotheses

To assess how much AI-generated texts affect the believability of a sender, we duplicate the research

design of Graefe et al. [7] and enhance it with fake news as the third category. Similar to Graefe et al. [7], we evaluate these three content categories regarding consumers' perceived credibility, readability, and journalistic expertise. Credibility relates to how accurate, trustworthy, fair, and reliable content is taken. Readability refers to how entertaining, interesting, vivid, and well-written content is seen. Journalistic expertise encompasses how coherent, concise, comprehensive, and descriptive content is understood. Based on the results of Graefe et al. [7] and the objectively false nature of fake news, negatively violating the expectations news consumers have about the content [24], we expect regarding credibility:

H1: Journalistic texts are more credible than fake news.

H2: AI-generated texts are as credible as journalistic texts.

H3: AI-generated texts are more credible than fake news.

Regarding readability, previous research has shown that readability of human-written texts is substantially higher than that of machine-written texts [20]. We expect fake news to be between these two types, as fake news is also human-written, but its origin and development does not hold up to journalistic standards [25]. Thus, we suggest:

H4: Journalistic texts are more readable than fake news.

H5: AI-generated texts are less readable than journalistic texts.

H6: AI-generated texts are less readable than fake news.

When it comes to journalistic expertise, machine authorship is usually perceived negatively [26]. Yet, this effect is not as pronounced when consumers read only one article, as is the case in this study [27]. In addition, when they are factual and accurate, texts written by a software are perceived as such [28]. In comparison, consumers often recognize the false or misleading nature of fake news and sharing rather occurs due to motivated reasoning [29], which is of little importance in the setting of this study. This leads to the following hypotheses:

H7: Journalistic texts have a higher level of perceived journalistic expertise than fake news.

H8: AI-generated texts have a similar level of perceived journalistic expertise as journalistic texts.

H9: AI-generated texts have a higher level of perceived journalistic expertise than fake news.

In addition to this manipulation of the believability of a source, i.e. examining texts that can be trusted and that cannot be trusted, we also look at trust on the consumer side. Here, we are interested to see how AI-generated text is perceived by consumers with higher

general trust in media compared to consumers with lower general trust in media. We base our study design on previous research [7, 8] and label content as AI-generated or do not label content at all, but additionally analyze consumer responses regarding credibility, readability, and journalistic expertise based on the level of media trust that participants exhibit.

The reasons for consumers' lacking trust in media are manifold, e.g., biases and heuristics [30]. For such participants, we assume that texts written by an AI seem more objective and less biased than those written by a journalist. Due to that, we assume that in all three ways of assessing content, i.e. perceived credibility, readability, and journalistic expertise, consumers' with little trust in media will tend more towards trusting an AI-generated text than consumers with high trust in media. Our respective hypotheses are:

H10: Consumers with low trust in media perceive the credibility of a text with an AI label higher than consumers with high trust in media.

H11: Consumers with low trust in media perceive the readability of a text with an AI label higher than consumers with high trust in media.

H12: Consumers with low trust in media perceive the journalistic expertise of a text with an AI label higher than consumers with high trust in media.

3. Method

To empirically test our hypotheses, we conducted an experiment in the form of an online survey. One of six different text variants was presented to each participant. They had to read the text and answer questions on their perception of the variables that are of interest for our study.

Three text designs were chosen for this experiment that differ as described above: a journalistic text from a quality newspaper, a conspiracy text containing fake news, and a text generated by an AI. Due to the experiment being conducted during the lockdown caused by the corona pandemic, we took advantage of the strong interaction of consumers with information on COVID-19 and the resulting abundance of available texts and selected articles about this topic. To avoid topic-specific difference [2], the topic of all texts was the same. The journalistic article was chosen according to scientific criteria [11] from the online portal of the *Süddeutsche Zeitung*, a renowned news media company in Germany. This article dealt with the typical symptoms of an infection with the coronavirus [31]. The conspiracy text also dealt with symptoms of COVID-19 and had been widely shared within Germany, particularly via the messaging service WhatsApp. It had been declared to be fake news by a consumer protection service [32]. The AI-generated text was an excerpt from

the daily corona newsletter of BR24, a channel of the German public broadcaster Bayerischer Rundfunk [33]. All texts were in German.

Before conducting the study, the selected texts were edited to reflect a similar appearance to eliminate possible effects of design elements, thus enabling the results to be comparable. The texts were also shortened to a length of approx. 500 words. This length was chosen to have enough space to give participants substantial insights as is usual for journalistic texts but reading would not take too much time to risk losing participants' attention. Information indicating the source of the text as well as illustrations and graphics were removed. Individual formatting of the text was also removed. This means that participants of the study were shown only pure texts with comparable length.

After that, two versions were created for each of the three texts. One version contained a manipulation of the byline marked as "Author: This text has been automatically generated using an AI". In this case the participants were correctly or incorrectly told that the text was created by an AI. The second version of the texts did not contain any labels and information about the author of the text.

To measure credibility, readability, and journalistic expertise, we used scales that have proven reliability in their application by Graefe et al. [7]. Our evaluation of media trust is based on a newly developed scale by Strömbäck et al. [34]. They have developed a framework that analyses media trust on different levels: general trust in news media, trust in the media type, trust in individual media brands, trust in journalists, and trust in media content. In the context of this work, the relationship between general media trust and the assessment of computer-generated news is analyzed. Thus, for the operationalization of general media trust, the proposed items for the analysis of general trust in news media by Strömbäck et al. [34] were used. The scale consists of five items, which were rated on a 6-point Likert scale. For the analysis, a general score of general media trust was calculated as the mean item values.

The actual experiment was conducted by a total of 277 participants obtained via a mixture of convenience and snowball sampling. We distributed the questionnaire via SoSci Survey to students of our institution in Germany. At the start of the survey, participants were told that they were to answer a survey regarding news on COVID-19. We did not let them know the true aim of our study at this point in time.

The first part of the study consisted of questions on sociodemographic and personality characteristics. In the second part, participants were assigned by the tool to one of the six experimental groups and read the respective text. In the third part, we asked questions

regarding the scales mentioned before. Questions include, e.g., how accurate, well written, or coherent the text presented to the participants was [7].

After finishing the study, participants that read the fake news text were debriefed by stating the true source of the text and telling them that its content was made up. The true aim of our study was also disclosed to all participants. They were then requested to give the link to the questionnaire to other people they know, but who belonged to an older generation and without telling these other people the true aim of our study. This way, we ensured that our sample covered different age and gender groups and different news reading habits. The link to the questionnaire was valid for 21 days. After that period, we considered the content of our texts of diminishing relevance in a journalistic context for our audience and stopped data collection.

		news text design		
		journalistic text	conspiracy text	automated computer-generated text
manipulation of byline text	without labels	n = 45	n = 44	n = 44
	with AI-label	n = 45	n = 46	n = 39

Figure 1. Cross-sectional design

Fourteen participants had to be excluded before data analysis in SPSS due to invalidity of their answers in the manipulation check regarding the content of the news text or due to obvious patterns in answering the questionnaire. This resulted in a sample size of n=263, distributed in the six experimental groups through automated randomization (see Figure 1) with roughly equal distribution according to age and gender. 167 of the participants were female and thus made up the majority of the participants (63.5%). 96 participants were male (36.5%) and none were diverse. Regarding age, 54 participants were between 18-23 years old (20.5%), 98 participants were 24-39 years old (37.3%), 61 participants 40-55 years old (23,2%) and 50 participants 55 years or older (19%).

4. Results

As a first step, the different text designs (journalistic text, fake news text, AI-generated text) were examined regarding differences in the assessment

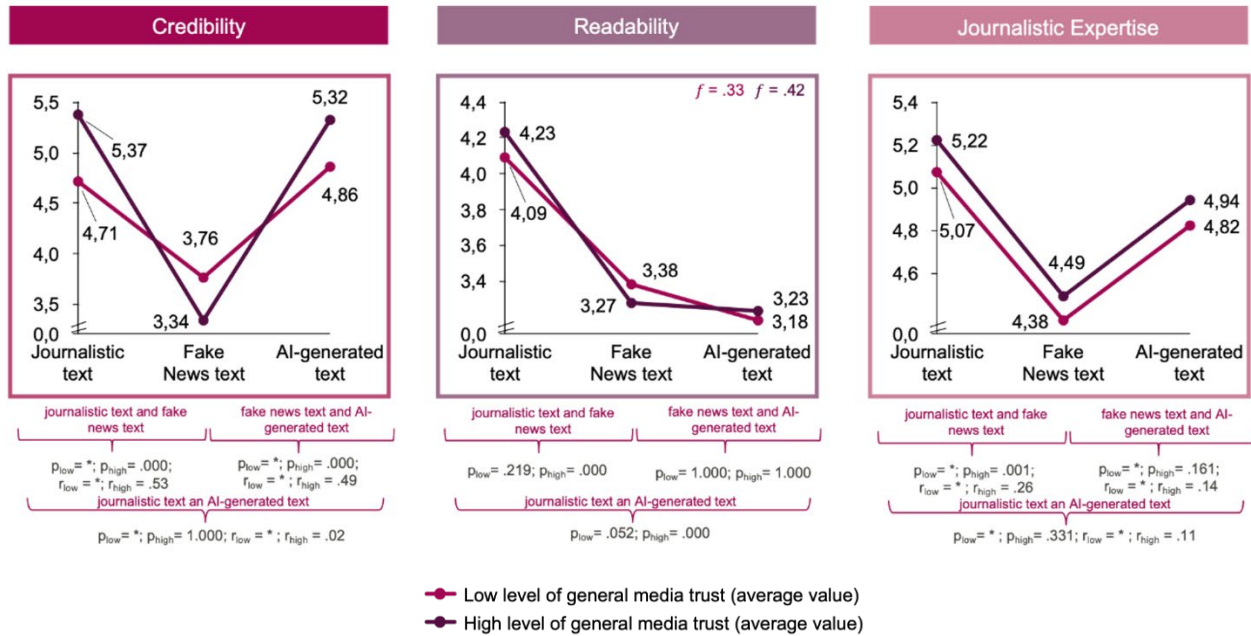


Figure 2. Participant assessment of the three text categories

* indicates that pairwise comparison is not significant

of credibility, readability, and journalistic expertise. Because the readability score was normally distributed according to a Kolmogorov-Smirnov test, univariate analysis of variance (ANOVA) was performed for this factor. For credibility and journalistic expertise, we applied Kruskal-Wallis tests. The results can be seen in Figure 2, where we additionally distinguished between the answers of participants with high (n=200, 76%) and low levels (n=63, 24%) of trust in media. It becomes apparent that regarding credibility, there is no significant difference between journalistic texts and AI-generated texts ($p > 0.05$). However, both of these kinds of texts are significantly more credible than the fake news text ($p < 0.001$). Thus, our first three hypotheses are supported.

Table 1. Hypotheses tests

Hypothesis	Result
H1: Journalistic texts are more credible than fake news.	Supported
H2: AI-generated texts are as credible as journalistic texts.	Supported
H3: AI-generated texts are more credible than fake news.	Supported
H4: Journalistic texts are more readable than fake news.	Supported

H5: AI-generated texts are less readable than journalistic texts.	Supported
H6: AI-generated texts are less readable than fake news.	Not supported
H7: Journalistic texts have a higher level of perceived journalistic expertise than fake news.	Supported
H8: AI-generated texts have a similar level of perceived journalistic expertise as journalistic texts.	Not supported
H9: AI-generated texts have a higher level of perceived journalistic expertise than fake news.	Not supported
H10: Consumers with low trust in media perceive the credibility of a text with an AI label higher than consumers with high trust in media.	Not supported
H11: Consumers with low trust in media perceive the readability of a text with an AI label higher than consumers with high trust in media.	Supported
H12: Consumers with low trust in media perceive the journalistic expertise of a text with an AI label higher than consumers with high trust in media.	Not supported

Regarding readability, our results show that there is a significant difference between journalistic texts and AI-generated texts as well as fake news ($p < 0.05$), but no significant difference between the latter two ($p > 0.05$). Thus, we found support for H4 and H5, but not for H6. AI-generated texts and fake news are perceived similarly bad when it comes to readability.

For journalistic expertise, we found significant differences between journalistic texts and fake news ($p < 0.01$), but no significant differences between journalistic texts and AI-generated texts or between fake news and AI-generated texts ($p > 0.05$). The visualization of the means in Figure 2 indicates that there might be a hierarchy between the three texts with journalistic texts at the top, followed by AI-generated texts and with fake news being at the bottom, but with our sample size these effects did not appear significant. Thus, we found support for H7, but not for H8 and H9. Table 1 summarizes the results of our hypotheses tests.

In this first evaluation of the manipulation of content sources (journalistic text, fake news, AI-generated text), no significant differences between participants with low and high trust in media could be found. The distribution of the mean values regarding credibility indicates that there might be a tendency for people with low trust in media to believe fake news more than people with high trust in media and to believe journalistic texts and AI-generated texts less, but our data analysis did not prove these effects to be significant.

Four our second data analysis, we looked at the effect of labeling texts as AI-generated or not labeling the texts at all for people with low and high trust in media (see Figure 3). For readability, a t-test provided a significant result. Participants with high general media

trust assessed a text with AI labeling worse ($M = 3.38$, $SD = 1.13$, $n = 93$) than a text without any labeling ($M = 3.75$, $SD = 1.21$, $n = 107$). Other effects for participants with high trust in media were on the edge of significance, e.g., $p = 0.052$ for journalistic expertise. However, as the results were not clearly within the accepted thresholds, we could not establish statistically significant differences between labeling texts as AI-generated or not for participants with high trust in media.

When it comes to participants with low trust in media, the picture is much clearer. For them, there is no significant difference between labeling a text as AI-generated or not. The texts are perceived similar in terms of credibility, readability, and journalistic expertise ($p > 0.05$). Regarding our hypotheses, this means that participants with low trust in media are indifferent to AI-generated texts, while participants with high trust in media perceive non-AI texts as more readable. While this supports our H11, we could not identify a significant relationship for H10 and 12.

5. Discussion and Implications

With the two analyses described above, we provide rich insights into how consumers perceive AI-generated texts, especially depending on their level of trust in media and in comparison to fake news. Our first analysis relates to the believability of the content, our second analysis to consumer trust in AI.

In the first analysis, our findings confirm previous research [4]. Journalistic texts and AI-generated texts are perceived similarly credible, but AI-generated texts are much less readable than journalistic texts. However, in contrast to previous research, we included fake news

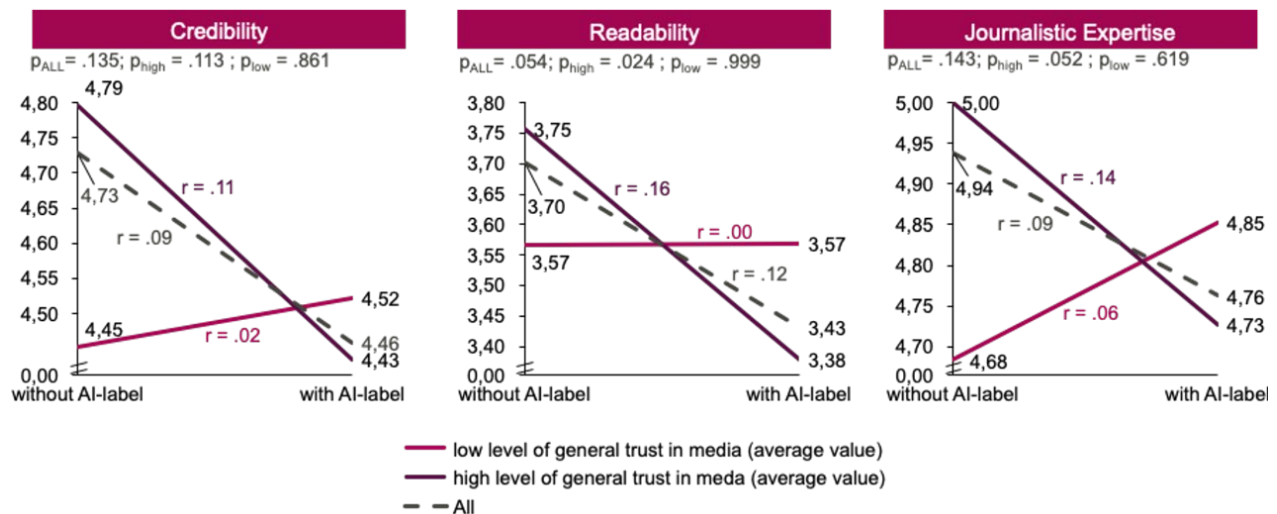


Figure 3. Effects of labeling texts as AI-generated for participants with low/high trust in media

as a third category in the analysis. This allowed us to evaluate how AI-generated content actually performs on the full scale of believable content, i.e. true journalistic content on the one side and objectively false fake news on the other side. As expected, AI-generated texts are more credible than fake news texts. However, the picture is not so clear when it comes to journalistic expertise, where we could find no significant differences to journalistic texts nor to fake news. The most surprising is our finding regarding readability, where AI-generated texts are not only significantly worse than journalistic texts, but on par with fake news. For media companies, this means that they have to pay attention especially on the aspects constituting readability, i.e. AI-generated texts need to be entertaining, interesting, vivid, and well-written. Improving these aspects is of paramount importance to lift media companies' AI-generated texts to a higher level of perceived readability and distinguish them clearly from fake news.

In the second analysis, we found that participants with high trust in media rated texts significantly worse in terms of readability when they were labeled as written by an AI-independent of whether that was true. Our data displayed similar tendencies for credibility and journalistic expertise, but these effects were not statistically significant, albeit borderline in the case of journalistic expertise. For participants with low trust in media, neither an effect nor a tendency towards human-written texts could be found. As mentioned, trust is a sensitive topic especially for news companies and providing AI-generated texts could decrease trust in media due to AI being non-transparent [6]. Yet, in our experiment labeling texts as produced by AI does not diminish credibility, readability, and journalistic expertise for many consumers, especially those with low trust in media. This finding especially enhances the results of Waddell [26], who attributed differences in perception caused by human or machine authorship to anthropomorphism and negative expectancy violations [24]. Our results indicate that besides these two attributions, consumers' general trust in media also influences perception of authorship.

Regarding the inclusion of fake news as a third category besides human-written and AI-generated texts, our study underscores that this additional category leads to more nuanced insights. Research has already shown that in terms of credibility, human-written and automatically created content are usually perceived rather similar [2], but differences in quality and especially readability occur, favoring human-written texts [20]. Through our findings, we now also know that readability of AI-based texts is not only perceived worse than that of human-written texts, but as bad as that of fake news, which can hurt the believability of a

messages' sender. With fake news as a third category, we might also place consumers' perception of journalistic expertise of AI-based texts between that of human-written texts and fake news, showing that the slight difference found in consumers' perception of quality in previous research [20] is also an improvement over the perception of fake news, but the significance of our findings does not allow for drawing this conclusion.

All in all, our two analyses emphasize that incorporating trust in research on consumer perception of AI-generated news is just as important as enhancing the previously used dichotomy of AI-generated and human-written texts with fake news as a third category. The insights of our study show that consumers' general trust in media matters when consuming AI-generated content. Furthermore, they demonstrate that fake news is an appropriate end of a content credibility scale, ranging from objectively well-written journalistic content to objectively false texts. Both of these contributions of our paper broaden the foundation for theory development in the area of automated news, where software is taking over tasks central for the identity and self-image of journalists and media companies with yet uncertain outcomes [35].

Finally, both analyses presented in this paper exhibit findings that are on the verge of significance. While this could mean that the effects are indeed not significant, the distribution of means displayed in the figures shows a clear trend. Other causes for such non-significant trends could be a too small sample size, bias due to participants with high levels of trust in media having a share of $\frac{3}{4}$ of our sample, or non-ideally chosen texts for the experiment. Especially the former must be considered, as we had to stop the experiment after 21 days to not risk polluting data due to outdated information. In any case, our findings highlight a need for more discussions and further studies on the interplay of trust and AI in media and communication.

6. Conclusion

Our empirical research intended to answer how trust in media influences consumers' perception of AI-generated media content. For this purpose, we first varied the believability of the content by using three different text designs on COVID-19. We then analyzed how consumer trust in media affects the perception of texts that were manipulated in their byline and (in-)correctly marked as AI-generated. The participants were randomly shown one of three selected texts in different formats either with or without manipulation of the byline, i.e. participants were divided in six groups in total. Our results highlight that AI-generated content is already comparable to journalistic content and superior to fake news in terms of credibility, but statistically little

different from fake news when it comes to readability. Furthermore, people with high and low trust in media perceive content labeled as AI-generated differently, independent from whether that content had really been written by an AI.

These insights expand previous research [e.g., 2, 7, 8, 20] by allowing AI-generated content to anchor on a scale ranging from true journalistic content to objectively false fake news. In addition, we show that consumer trust in media needs to be considered when texts labeled as AI-generated are offered [26], as the perception especially of readability depends on the general level of trust consumers have in media.

7. References

- [1] Balakrishnan, T., M. Chui, B. Hall, and N. Henke, *Global Survey: The State of AI in 2020*, McKinsey & Company, New York, 2020.
- [2] Wölker, A., and T.E. Powell, “Algorithms in the Newsroom? News Readers’ Perceived Credibility and Selection of Automated Journalism”, *Journalism*, 22(1), 2021, pp. 86-103.
- [3] Thurman, N., S.C. Lewis, and J. Kunert, “Algorithms, Automation, and News”, *Digital Journalism*, 7(8), 2019, pp. 980-992.
- [4] Crawford, K., R. Dobbe, T. Dryer, G. Fried, B. Green, E. Kaziunas, A. Kak, V. Mathur, E. McElroy, A. Nill Sánchez, D. Raji, J.L. Rankin, R. Richardson, J. Schultz, S. Myers West, and M. Whittaker, *AI Now 2019 Report*, AI Now Institute, New York, 2019.
- [5] J.G. Saliby, “Survey on Natural Language Generation”, *International Journal of Trend in Scientific Research and Development*, 3(3), 2019, pp. 618–622.
- [6] E. Reiter, and R. Dale, “Building Applied Natural Language Generation Systems”, *Natural Language Engineering*, 3(1), 1997, 57–87.
- [7] A. Graefe, M. Haim, B. Haarmann, and H.-B. Brosius “Readers’ Perception of Computer-generated News: Credibility, Expertise, and Readability”, *Journalism*, 19(5), 2018, pp. 595–610.
- [8] H.A.J. van der Kaa, and E.J. Kraemer, “Journalist versus News Consumer: The Perceived Credibility of Machine Written News”, *Proceedings of the Computation + Journalism Symposium*, New York, 2014.
- [9] M. Riveiro, and S. Thill, “That’s (not) the Output I Expected! On the Role of End User Expectations in Creating Explanations of AI Systems”, *Artificial Intelligence*, 298, 2021, article 103507.
- [10] van der Linden, S., C. Panagopoulos, and J. Roozenbeek, “You are Fake News: Political Bias in Perceptions of Fake News”, *Media, Culture & Society*, 42(3), 2020, pp. 460-470.
- [11] Bachmann, P., M. Eisenegger, and D. Ingenhoff, “Defining and Measuring News Media Quality: Comparing the Content Perspective and the Audience Perspective”, *The International Journal of Press/Politics*, 2021, DOI: 10.1177/1940161221999666.
- [12] Lazer, D.M.J., M.A. Baum, Y. Benkler, A.J. Berinsky, K.M. Greenhill, F. Menczer, M.J. Metzger, B. Nyhan, G. Pennycook, D. Rothschild, M. Schudson, S.A. Sloman, C.R. Sunstein, E.A. Thorson, D.J. Watts, and J.L. Zittrain, “The Science of Fake News”, *Science*, 359(6380), pp. 1094-1096.
- [13] A. Gatt, and E. Kraemer, “Survey of the State of the Art in Natural Language Generation: Core Tasks, Applications and Evaluation”, *Journal of AI Research*, 61, 2017, pp. 75-170.
- [14] Gramatke, M., J. Gegner, S. Bauer, M. Boenisch, V.D.B. Le, *Cognitive Artificial Intelligence: The Invisible Invasion of the Media Business*. Report, Deloitte, undated.
- [15] Hansen, M., M. Roca-Sales, J. Keegan, and G. King, *Artificial Intelligence: Practice and Implications for Journalism*, Tow Center for Digital Journalism, Brown Institute for Media Innovation, 2017.
- [16] C. Clerwall, “Enter the Robot Journalist: Users’ Perception of Automated Content”, *Journalism Practice*, 8(5), 2014, pp. 519–531.
- [17] J. Newhagen, and C. Nass, “Differential Criteria for Evaluating Credibility of Newspapers and TV News”, *Journalism Quarterly*, 66(2), 1989, pp. 277–284.
- [18] A. Miller, and D. Kurpius (2010). “A Citizen-Eye View of Television News Source Credibility”, *American Behavioral Scientist*, 54(2), 2010, pp. 137–156.
- [19] Adam, M., M. Wessel, and A. Benlian, “AI-based Chatbots in Customer Service and their Effects on User Compliance”, *Electronic Markets*, 31, 2021, pp. 427–445.
- [20] Graefe, A., and Bohlken, N., “Automated Journalism: A Meta-Analysis of Readers’ Perceptions of Human-Written in Comparison to Automated News”, *Media and Communication*, 8(3), 2020, 50-59.
- [21] R.C. Mayer, J.H. Davis, and F.D. Schoorman, “An Integrative Model of Organizational Trust”, *Academy of Management Review*, 20(3), 1995, pp. 709-734.
- [22] M.J. Metzger, and A.J. Flanagin, “Credibility and Trust of Information in Online Environments: The Use of Cognitive Heuristics”, *Journal of Pragmatics*, 59(B), 2013, pp. 210-220.
- [23] T.-T. Lee, “Why They don’t Trust the Media: An Examination of Factors Predicting Trust”, *American Behavioral Scientist*, 54(1), 2010, pp. 8-21.
- [24] Burgoon, J.K., “A Communication Model of Personal Space Violations: Explication and an Initial Test”, *Human Communication Research*, 4(2), 1978, pp. 129-142.
- [25] Jang, S.M., T. Geng, J.-Y.Q. Li, R. Xia, C.-T. Huang, H. Kim, J. Tang, “A Computational Approach for Examining the Roots and Spreading Patterns of Fake News: Evolution Tree Analysis”, *Computers in Human Behavior*, 84, 2018, pp. 103-113.
- [26] Waddell, T.F., “A Robot Wrote This?”, *Digital Journalism*, 6(2), 2018, 236-255.
- [27] Haim, M., and Graefe, A., “Automated News”, *Digital Journalism*, 5(8), 2017, pp. 1044-1059.
- [28] Oh, C., J. Choi, S. Lee, S. Park, D. Kim, J. Song, D. Kim, J. Lee, and B. Suh, “Understanding User Perception of Automated News Generation System”, *Proceedings of*

- the 2020 CHI Conference on Human Factors in Computing Systems, Honolulu, April 2020.
- [29] Tsang, S. J., “Motivated Fake News Perception: The Impact of News Sources and Policy Support on Audiences’ Assessment of News Fakeness”, *Journalism & Mass Communication Quarterly*, 2020, DOI: 10.1177/1077699020952129.
- [30] Newman, N., and R. Fletcher, *Bias, Bullshit and Lies: Audience Perspectives on Low Trust in the Media*, Digital News Project, Reuters Institute for the Study of Journalism, University of Oxford, 2017.
- [31] Bartens, W., “Die typischen Symptome des Coronavirus [The Typical Symptoms of the Coronavirus]”, *Süddeutsche Zeitung*, 2020, <https://www.sueddeutsche.de/gesundheit/coronavirus-symptome-verlauf-covid-19-1.4851200>.
- [32] Sommer, K., “WhatsApp Kettenbrief: Ein italienischer Arzt, der im Shenzhener Krankenhaus in China arbeitete ... Echt oder Fake? [WhatsApp Chain Letter: An Italian Doctor Who Worked in the Shenzhen Hospital in China ... Real or Fake?]”, *Verbraucherschutz.com*, 2020, <https://www.verbraucherschutz.com/warnungsticker/watsapp-kettenbrief-mit-arztipp-wichtige-hinweise-oder-falschinformation>.
- [33] BR24, “Coronavirus: Aktuelle Zahlen für Bayern und Deutschland [Coronavirus: Current Figures for Bavaria and Germany]”, BR24, 2020, <https://interaktiv.br.de/corona-newsletter>.
- [34] J. Strömbäck, Y. Tsfati, H. Boomgaarden, A. Damstra, E. Lindgren, R. Vliegthart, and T. Lindholm, “News Media Trust and its Impact on Media Use: Toward a Framework for Future Research”, *Annals of the International Communication Association*, 44(2), 2020, pp. 139–156.
- [35] Lewis, S.C., A.L. Guzman, T.R. Schmidt, “Automation, Journalism, and Human–Machine Communication: Rethinking Roles and Relationships of Humans and Machines in News”, *Digital Journalism*, 7(4), pp. 409–427.