Machine Credibility: How News Readers Evaluate AI-Generated Content

Abstract

The advent of AI-generated news as a novel form of content demands renewed attention toward modes of understanding reader perceptions. This research sought to answer: What evaluative criteria do readers use to perceive automated news content? To answer this, the study employed a two-phase survey methodology designed to elicit reader perceptions of AI-generated news. Phase 1 yielded 26 dynamic descriptor words and reflected broad social perceptions of AI. In Phase 2, a series of exploratory factor analyses (EFA) was conducted on results of a survey using the 26 items obtained in phase 1 to uncover underlying factors contributing to differences in how readers ranked articles based on the aforementioned descriptor words. In both phases, readers were informed at the beginning of the survey that the articles were generated using AI. The first set of exploratory factor analysis results were obtained using varimax rotation, which revealed five salient factors underlying the 26 descriptors labeled Quality, Engagement, Alienation, Effort, and Coherence. The second exploratory factor analysis used oblimin rotation, which contrastingly revealed nine salient factors, which were labeled Credibility, Prolixity, Engagement, Substance, Clarity, Alienation, Complexity, Effort, and Neutrality. When compared with the results of factor analyses for human-generated news content, the findings offer new constellations of terms that reflect the dimensions that readers attend to in articles attributed to artificial intelligence.

Keywords: Automation, Journalism, Generative, Perceptions, Quality, Survey

Introduction

The emergence of algorithmic integration in newsrooms, manifested in part in the form of article writing, has gained significant momentum over the past decade (Biswal and Kulkarni, 2024). This acceleration is largely powered by the advent of advanced large language models (LLMs). The public release of LLMs such as OpenAI's ChatGPT and Google's Bard garnered significant hype and media attention; such advancements underscore the escalating relevance of generative AI (genAI) across numerous domains (Espejel et al., 2023). Parallel to this technological evolution, the academic world has seen a surge in research exploring reader perceptions of AI-generated news (see, e.g., Graefe and Bohlken, 2020; Henestrosa et al., 2023). AI can generate textual content for news articles in two major ways: automated news writing and natural language processing. With automated news writing, AI systems can write news articles based on structured data inputs such as financial reports, sports scores, weather updates, or election results (see, e.g., Associated Press News, 2021; Kunert, 2020). Algorithms analyze the data and generate coherent articles, often in real-time. The other form of AI news generation, natural language processing (NLP), uses technologies that enable AI to understand and generate human language, making it possible to produce articles that mimic human writing styles and tones (Dergaa et al., 2023). Previous academic explorations of AI-generated news perceptions have focused mainly on classic journalistic metrics such as credibility, quality, and readability, all key to human-authored news (Sundar, 1999; Graefe and Bohlken, 2020). However, the introduction of non-human authorship requires a reassessment of these well-established evaluative metrics (Panagia, 2021). The purpose of this research, therefore, is to probe deeper into the criteria readers employ when interacting with automated news content.

Prior explorations by media studies scholars into the perceptions of print and online news among receivers have employed exploratory factor analysis (EFA) to delve into the complex interplay of various descriptors used by readers to characterize news content (Sundar, 1999). This approach is effective in its ability to reveal the underlying factor structure of these descriptors, offering a nuanced understanding of how news, both in traditional print and emerging online formats, is evaluated by its consumers. In this context, EFA serves as a robust analytical tool capable of sifting through multifaceted survey data to uncover latent constructs or dimensions of reader perceptions (Fabrigar et al., 1999). By identifying these constructs, this exploratory research method offers invaluable insights into the factors that most significantly impact how news content is received and judged, and how the integration of AI in journalism presents new challenges and opportunities.

In the broader context of news research, much emphasis has been placed on distinguishing news content from other communication forms (Robinson and Levy, 1986; Sundar, 1999; Harcup and O'Neill, 2017). This is based on the understanding that each content category has its unique evaluative criteria considered pivotal. However, the advent of AI-generated news as a novel content concept demands renewed attention toward modes of understanding reader perceptions. To make sense of how readers perceive AI-generated news, it is of paramount importance to unravel the criteria that underpin these perceptions. As the media world assesses a new era of journalism, this research becomes crucial to understanding and navigating a dynamic landscape.

Building upon the foundational works in journalism research, this study of AI-generated news perceptions is particularly timely. An evaluation of the perceptions of AI-generated news content aims foremost to understand essential components of a good AI-written news story. In order to understand this, it is crucial to assess the essential components of a good news story, and whether or not these components are the same as those of a good AI news story; and furthermore, should they be the same? Moreover, are these components attended to by receivers of this content (i.e. news readers)? This study considers these factors as its foundation in its exploration of the evaluative criteria readers use to perceive automated news content.

Literature Review

The impact of generative AI on journalism has been transformative, redefining the scope and methodology of news production. The integration of AI tools in journalism is not just a futuristic concept but a present reality, with applications ranging from routine reporting to complex data analysis. One of the most straightforward applications is in sports journalism, such as automating reports on high school football scores. Similarly, AI has been instrumental in analyzing vast troves of leaked documents, sifting through extensive data to uncover critical insights. These tasks, traditionally time-consuming and labor-intensive, are now more efficient and accurate with AI assistance (Beckett, 2023).

The COVID-19 pandemic serves as a contemporary case study of news automation capabilities. As noted by Danzon-Chambaud (2022), the outbreak saw governments and health authorities release substantial amounts of open-source data, accessible through structured datasets or APIs. ChatGPT, with its advanced language generation capabilities, presents another intriguing case study for the use of NLP in news automation. As Diakopoulos (2023) points out, while automated writing based on structured data is a long-standing practice in journalism, models like ChatGPT bring a nuanced complexity to this process. ChatGPT can produce fluently written text from structured data inputs, though it does involve statistical sampling and may not perform mathematical operations.

Automated writing from unstructured data is more challenging than the structured data case. Generative AI models can extract structured data from unstructured sources, such as press releases, and then use this to generate coherent text (Bandi et al., 2023). For instance, a model might extract details like dates, times, locations, and descriptions of community events from a press release. This structured data is then fed back into the model to produce a written account of the event. Ultimately, the burgeoning influence of generative AI in journalism inevitably brings to the forefront the critical importance of understanding how readers perceive content generated by these advanced tools.

News organizations' decisions to deploy generative AI tools for content creation are significantly influenced by considerations of how such content is perceived by their audience. The criteria traditionally used to evaluate news value, derived primarily from research on human-written content, may require reconsideration. In their landmark 1965 study, one of the first to examine these criteria, Johan Galtung and Mari Ruge asked "How do 'events' become 'news'?", leading them to identify twelve key factors that play a pivotal role in this transformation (Galtung & Ruge, 1965). Their research focused on how overseas events were reported as foreign news in the Norwegian press. These factors, ranging from frequency and threshold to negativity and reference to elite nations and people, provided a framework for understanding the selection and distortion processes in news reporting.

In the decades following Galtung and Ruge's study, the academic discourse surrounding news selection and presentation has continued to evolve. Joye et al. (2016) note that while some studies have confirmed Galtung and Ruge's findings, others have raised methodological concerns and proposed additional news factors, advocating for a more nuanced and expanded model of news selection. For example, Sundar (1999) ran a study that delved into the critical criteria used by news consumers in their perceptions of news by using factor analyses of inter-correlations among measurements. These analyses yielded four key factors: Quality, Credibility, Liking, and Representativeness. Joye and colleagues highlight three critical areas for contemporary news value research: assessing the relevance of Galtung and Ruge's hypotheses in the context of today's data-rich environment, integrating the changing societal and cultural contexts in news selection, production, and reception, and aligning the study of news values with the realities of global journalism.

Davide Panagia (2023) describes technologies such as ChatGPT as "systems that barter in Bayesian probabilities rather than mimetic representations" -- and thus, the frameworks we use to critique or critically think about these technologies are obsolete insofar as "we think of the activity of 'challenging' as a critical operation that negates a representation or an identity" (p. 2). An approach to critically evaluating media that accounts for this distinction would thus benefit from an exploratory understanding of how the intended audience of such media perceives it. If news media is no longer a representation of mind, but rather an accumulation of probabilistic calculations, readers are dealing with a different entity of media that is nonetheless presented as possessing those same intrinsic qualities. Moreover, if readers are cognizant of the AI authorship of content, their own evaluative criteria might shift, necessitating the development of new metrics within academic research to aptly compare human and AI-generated news.

For example, the news values identified by Galtung and Ruge in 1965, or Sundar in 1999, may be less relevant in terms of their prioritization by news consumers in the context of AI-generated content. This difference in importance to consumers is independent of how they are traditionally prioritized for marketing and administrative purposes in the newsroom. Existing research on reader perceptions has relied on Likert-type or semantic differential scales that use adjectives deemed relevant by researchers and force receivers to rate news articles and sources along the dimensions researchers propose (Graefe and Bohlken, 2020; Wang and Huang, 2024). However, the relevant psychological dimension(s) along which participants vary in response to stimuli may be different than the adjectives deemed relevant by researchers given the potential differences between AI-generated content and human-generated content in terms of the factors to which readers attend. These dimensions could shift for two reasons: 1) if the content or message itself is different, but also 2) if the content is explicitly attributed to a different source.

In considering the evolving landscape of news media, it's crucial to distinguish between news content and other content types for both media industry and administrative purposes. However, an intriguing question arises when we consider the psychological processing of this content by audiences. Does the categorical distinction between news and other forms of content significantly impact how readers perceive and process the information presented to them? This question invites a deeper exploration of the psychological underpinnings that govern media consumption. Similarly the distinction between AI-generated and human-authored news is a critical consideration from a media production standpoint. Does the explicit categorical distinction between AI and human authorship impact how readers perceive and process information? Yet, from a consumer perspective, the importance of this distinction may not be as pronounced. Readers' engagement with news content, their trust in its credibility, and the value they ascribe to it could be influenced more by the content's inherent qualities — such as accuracy, relevance, and comprehensiveness — rather than by knowledge of its authorship.

Departing from the canon of empirical studies that have attempted to isolate the impacts of both the content itself and the attribution of authorship, this study seeks to dive deeper into the constructs underlying the metrics used in these studies themselves. Are metrics like credibility still relevant, and also, should they still be relevant? The way people describe news content that is explicitly written by AI will potentially illuminate new constructs and constellations of concepts brought on by novel content and source cues.

Methodology

This research sought to answer:

RQ: How do readers identify and evaluate the presence of potential hidden human manipulation in AI-generated news content?

In light of the nuanced and sometimes contradictory perceptions that readers have towards automated news, this research seeks to delve deeper into the evaluative criteria employed by readers when interacting with AI-generated content. Understanding these criteria is vital for news organizations as they navigate the complex dynamics of integrating AI into their journalistic practices, balancing technological efficiencies with reader trust and acceptance. Thus, the central research question aims to uncover the specific factors that influence reader perceptions, offering insights into how automated news is received. The answers to this question are crucial for shaping the future deployment of generative AI tools in journalism, ensuring that the content produced not only leverages the advancements in AI technology but also resonates positively with intended audiences.

To answer the RQ, the study employed a two-phase survey methodology designed to elicit reader perceptions of AI-generated news. Phase 1 of the survey sought to gain a qualitative understanding of how people describe automated news content. After obtaining long-form impressions from respondents, individual terms were isolated and used as variables for Phase 2. Reader rankings of articles based on said variables formed the basis for the exploratory factor analysis, in which connections between variables were identified to extract factors underlying reader perceptions.
The methodological choices for the two-phase survey methodology in this study were strategically designed to capture a comprehensive understanding of reader perceptions of AI-generated news. In Phase 1, the focus on qualitative, long-form survey responses was crucial for gathering in-depth, descriptive insights from readers. This approach allowed participants to freely express their impressions and thoughts, providing a breadth of perspectives on AI-generated content. The qualitative nature of this phase was instrumental in capturing the nuanced and varied ways in which readers perceive and articulate their experiences with automated news, beyond the constraints of predefined response options.

Phase 2, building on the groundwork laid in Phase 1, involved a more structured approach with a questionnaire eliciting Likert-scale rankings. This phase was designed to quantify the descriptors obtained from Phase 1, allowing for a systematic, measurable comparison of reader perceptions across different AI-generated articles. The choice to conduct an exploratory factor analysis on the correlations between these ranked variables was a deliberate one. It enabled the identification of underlying factors that influence reader perceptions, moving beyond individual descriptors to reveal broader patterns and themes in how AI-generated news is evaluated.

Note: the study has been deemed exempt from review by the UCLA IRB ethics committee (IRB#22-000989)

Survey Design: Phase 1
In Phase 1, participants were asked to detail their impressions and thoughts regarding an AI-written news article using descriptive adjectives. Specifically, the survey asked each participant to "List the thoughts that come to your mind after reading the article" and "List 2-10 adjectives describing the article" This phase was first designed and piloted in March 2023 on Amazon Mechanical Turk (MTurk), an online platform widely used for distributing surveys and tasks that require human intelligence. For the pilot survey, responses were gathered from 50 MTurk workers. Following a successful pilot, the final version of the survey was designed and distributed through YouGov, an international research data and analytics group, in April 2023. A larger sample size for this stage was sought, resulting in responses from 100 participants. The article was generated using ChatGPT (GPT-4) based on an existing, human-written article in Reuters, an international news service, titled "Internet Archive's digital book lending violates copyrights, US judge rules". To generate

the article, ChatGPT was given the prompt "Write an article in the style of Reuters with the title: 'Internet Archive's digital book lending violates copyrights, US judge rules'".

In both Phase 1 and Phase 2, survey participants were selected using the recruitment protocols of Amazon Mechanical Turk. Workers were required to be Masters to complete each task. The Masters qualification is assigned to workers when they have demonstrated superior performance over a period of time across thousands of Human Intelligence Tasks (HITs). Additionally, workers were required to have an HIT Approval Rate greater than 95%. For Phase 1, each worker was paid \$2.50 to complete the survey. Demographics of workers on MTurk have been found to mirror the U.S. population: majority female by a slight margin, with approximately three-quarters identifying as white. The overall distribution of household income among MTurk workers has been found to be within a few percentage points of the U.S. population for each income bracket except for households making more than \$150,000 per year (Moss et. al., 2023).

Survey Design: Phase 2

The survey for Phase 2 was designed on Qualtrics and deployed on Amazon Mechanical Turk. The recruitment process for Phase 2 mirrored that of Phase 1. For Phase 2, each worker was paid \$2.50 to complete the survey. A pilot survey for Phase 2 was conducted in June 2023 with a sample of 100 respondents. The pilot study allowed for the testing of the design and format of their survey, including the wording and order of questions, the layout of the survey, and the functionality of interactive elements. Following the pilot, an anonymous survey was carried out between August and October 2023 as part of the main Phase 2 study. For this survey, a sample of 261 respondents was obtained. This number was deemed adequate given the 26 variables being measured in the data set, and the general rule of thumb for adequate EFA sample sizes is 10 subjects per variable (Nunnally, 1978). Upon commencing the survey, each participant was shown three articles. To write each of the articles, ChatGPT (GPT-4), a large language model developed by OpenAI, was prompted with the title and source of an existing, human-written article with the same title. The output was not edited in any way. Participants were informed that each of the articles were generated using ChatGPT.

Each of the articles covered a different topic to provide a breadth of content for measurement: politics, finance, and technology. The human-written articles were sourced from international news outlet Reuters. For the AI-written articles, for each article title, ChatGPT was given the prompt "Write an article in the style of Reuters with the title: ___"After each article, participants were prompted with these instructions: In the following question, for each word, please rate how well the word describes the article above, from "describes very poorly" (1) to "describes very well" (5). At the end of the survey, each participant was given a unique survey completion code to submit via MTurk. After data collection was finalized, exploratory factor analyses (EFA) were conducted to scrutinize the differences in factor structures, thereby shedding light on potential variations in perceptions of stimuli.

Analysis: Exploratory Factor Analysis (EFA)

In the exploratory factor analysis of this study, two distinct factor rotation methods were employed: varimax and oblimin. The decision to include results from both varimax and oblimin rotations in the study was driven by the objective to gain a comprehensive understanding of the data. While orthogonal rotations like varimax simplify the structure and interpretation of factors, they might not always represent the true underlying relationships in the data, especially when factors are correlated. Oblique rotations like oblimin, although potentially more complex to interpret, can offer a more realistic picture of these inter-factor relationships. By examining the results of both rotation methods, the study leveraged the clarity and simplicity of orthogonal rotations and the realistic representation of factor correlations provided by oblique rotations, thereby ensuring that the findings were both interpretable and closely aligned with the structure of the data itself.

Results and Discussion

Phase 1

After collecting the data for Phase 1 of the study, the responses to the "List 2-10 adjectives describing the article" question were filtered and cleaned (reduced to individual adjectives) to ensure usability and reliability. This process resulted in the extraction of 44 unique descriptive words (59 total) used by respondents to characterize the AI-generated news content. Subsequent refinement and consolidation yielded 26 unique descriptors for the final list. Of the 44 descriptors, some were synonymous (e.g. Amazing and Fabulous) and as such were redundant for use in the exploratory factor analysis. Furthermore, certain measures were eliminated from the study because they are not appropriate as descriptors of all news, only certain news content or specific sources of news. Two other minor changes were made: first, the descriptor Flowing was changed to Coherent, a synonym, because the latter is a more widely understood term. Second, words with negative/opposite force were changed to their positive counterparts (Unbiased to Biased, Inconclusive to Conclusive) for the sake of clarity. The full list of descriptors used (bolded) along with the rationale for filtering given words (unbolded) is shown in the Appendix.

The final list of 26 words is shown below:

1. Amazing
2. Believable
3. Biased
4. Boring
5. Coherent
6. Complex
7. Concise
8. Conclusive
9. Detailed
10. Educational
11. Factual
12. Fair
13. Honest
14. Informative
15. Interesting
16. Long
17. Moral
18. Neutral
19. Precise
20. Rote
21. Technical
22. Thorough
23. Thought-provoking
24. Timely
25. Weird
26. Wordy

These descriptors encompassed traditional news story attributes such as fairness and neutrality but also included novel descriptors like "Technical," "Precise," and "Weird." This suggests that broader social perceptions of AI may influence article descriptions.

Phase 2

As previously mentioned, 261 responses were recorded for the Phase 2 survey. Each response included ratings of the 26 variables for three articles, for a total of 783 observations for each variable. Of the 20358 possible values in the data set, 35 were missing (NA). The analysis for

Phase 2 was performed using R and RStudio, version 1.1.42, which are widely recognized for their robust statistical capabilities and flexibility in handling complex data sets. The correlation matrix was analyzed using the Pearson method, known for its efficacy in measuring the linear correlation between variables. For the estimation method, this study utilized the minimum residual method, or minres. The method of estimating communalities, essential for understanding the shared variance in observed variables, involved using both h^2 and u^2 methods.

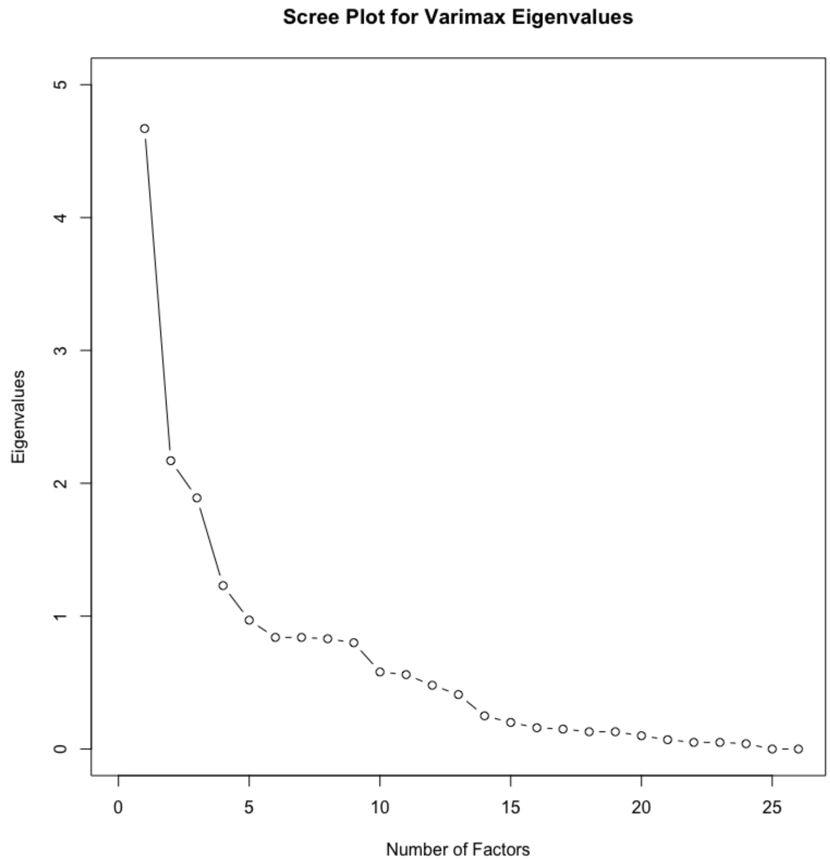
Before proceeding with the Exploratory Factor Analysis (EFA), it was crucial to determine the suitability of the collected data for such statistical processing. Two key tests were employed for this purpose: Bartlett's Test of Sphericity and the Kaiser-Meyer-Olkin (KMO) Measure of Sampling Adequacy. The Chi-square test statistic from Bartlett's test was 9189.899, with a p-value significantly less than 2.22 * 10^16, essentially approaching zero. This extremely low p-value strongly rejects the null hypothesis of the test, suggesting that the variables are related and, therefore, suitable for factor analysis. The overall KMO Measure of Sampling Adequacy was 0.92, far exceeding the minimum acceptable level of 0.50. A KMO value of 0.92 is considered excellent, indicating that a significant amount of variance might be explained by underlying factors and that the data is very suitable for EFA. In sum, both tests strongly indicated that the data was appropriate for Exploratory Factor Analysis, ensuring the reliability and validity of the subsequent analysis.

Exploratory Factor Analysis

In this study, the criterion for retaining factors was primarily based on eigenvalues, with additional consideration given to scree plot analyses (see Figs. 1 and 2 below). Factors with eigenvalues over 1 are typically considered significant as they explain more variance than a single observed variable. The sum of the squared loadings of each variable with a given factor (the column sum of the squared loadings matrix), which equal the factor's eigenvalue, were calculated in R using the fa() function. For the varimax rotation, the eigenvalues for the highest five factors were 4.67, 2.17, 1.89, 1.23, and 0.97. The fifth factor, with an eigenvalue slightly below 1, was also retained based on its proximity to 1 and the insights provided by the scree plot. Including the fifth factor was deemed important as it appeared to contribute meaningful information about the data structure based on its positioning on the scree plot.

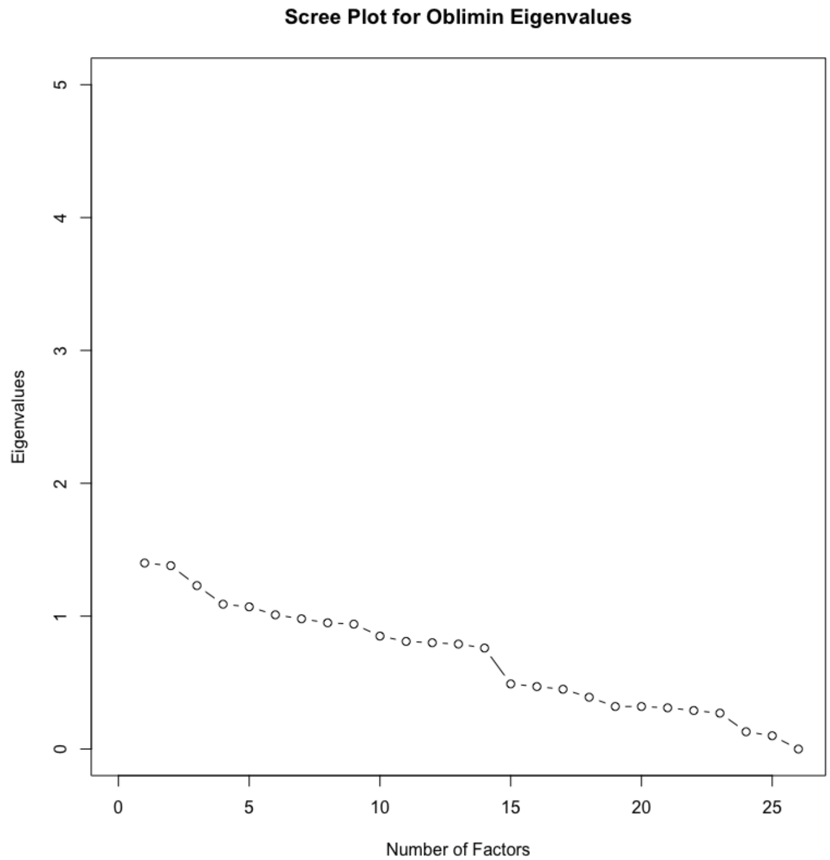
Figure 1

Scree Plot (Varimax Rotation)



In the oblimin rotation, which allows for correlation among factors, a total of nine factors were retained. The eigenvalues for these factors were 1.40, 1.38, 1.23, 1.09, 1.07, 1.01, 0.98, 0.95, and 0.94. Here, factors with eigenvalues close to but less than 1 were also included, as indicated by the scree plot analysis. This was based on the understanding that in oblique rotations, lower eigenvalues can still be meaningful due to the potential correlations between factors. Both eigenvalue criteria and scree plot analyses were instrumental in determining the number of factors to retain for each rotation method.

Figure 2
Scree Plot (Oblimin Rotation)



Tables 1 and 2 below summarize the eigenvalues for the factors retained under each analysis, including the proportion of variance and cumulative variance explained by each factor. Analysis of the table for the varimax rotation with the given eigenvalues indicates that the first factor explains a significantly larger portion of the variance in the data compared to subsequent factors. The high eigenvalue in the first factor suggests that this factor in particular is the most influential in explaining the variability in the dataset. The first factor accounts for 27% of the variance, which is a substantial amount, indicating that this factor captures a significant portion of the information in the data set. The following factors, with proportions of 12%, 11%, 7%, and 6%, contribute progressively less to the total variance, but each still represents a meaningful aspect of the data. These five factors together explain 62% of the cumulative variance in the data. While this variability explained value is a substantial amount in social science research, it also illustrates that there is a moderate amount of variance in the data that is not captured by these factors, indicating other considerations might be influencing variance in the data.

Table 1

Eigenvalues and Variance Explained (Varimax Rotation)

Factor Number	**Eigenvalues**	**Proportion of Variance**	**Cumulative Variance**

1	4.67	0.27	0.27
2	2.17	0.12	0.39
3	1.89	0.11	0.50
4	1.23	0.07	0.57
5	0.97	0.06	0.62

The eigenvalues obtained after the oblimin rotation range from 0.94 to 1.40 for the nine factors retained. The fact that these eigenvalues are relatively close to each other suggests that each factor contributes somewhat similarly to explaining the variance in the data, unlike in the varimax rotation. Each factor contributes between 5% to 8% of the variance. This even distribution further reinforces that the data's underlying structure contains no single factor dominating the explanation of variance. The 0.58 cumulative variance explained by the nine factors is significant, but similar to the value in the varimax rotation, it also shows that 42% of the variance in the data is not explained by these factors.

Table 2

Eigenvalues and Variance Explained (Oblimin Rotation)

Factor Number	**Eigenvalues**	**Proportion of Variance**	**Cumulative Variance**
1	1.40	0.08	0.08
2	1.38	0.08	0.16
3	1.23	0.07	0.23
4	1.09	0.06	0.29
5	1.07	0.06	0.35
6	1.01	0.06	0.42
7	0.98	0.06	0.48
8	0.95	0.05	0.53

9	0.94	0.05	0.58

Below, in tables 3 and 4, are the factor loadings in tabular format, along with the variables that load onto each salient factor, and labels for each salient factor. Salience was defined as a factor having loadings of above 0.30 from variables with their highest loading on said factor.

Table 3
Factor Loadings of News Stories Ratings - Varimax Rotation

Measure	Factor 1: Quality	Factor 2: Engagement	Factor 3: Alienation	Factor 4: Effort	Factor 5: Coherence*
Honest	**0.83**	0.16	-0.09	0.00	-0.05
Factual	**0.74**	0.16	-0.10	0.08	0.01
Fair	**0.73**	0.04	-0.15	-0.07	0.13
Believable	**0.68**	0.09	-0.21	0.00	0.15
Precise	**0.68**	0.34	-0.01	0.10	0.07
Thorough	**0.65**	0.30	-0.05	0.33	0.03
Informative	**0.62**	0.36	-0.19	0.25	0.04
Conclusive	**0.56**	0.31	0.04	0.13	-0.01
Concise	**0.53**	0.30	0.12	-0.19	0.21
Neutral	**0.52**	-0.08	0.06	-0.09	0.11
Detailed	**0.50**	0.34	-0.04	0.39	0.10
Moral	**0.50**	0.16	0.25	-0.07	0.00
Educational	**0.49**	0.39	0.04	0.22	0.05
Timely	**0.45**	0.28	-0.02	0.00	0.28
Interesting	0.32	**0.74**	0.06	-0.02	0.08
Thought-provoking	0.21	**0.62**	0.15	0.17	0.07
Amazing	0.17	**0.54**	0.35	0.20	-0.19

Weird	-0.23	0.06	**0.73**	0.10	-0.15
Rote	0.04	0.03	**0.60**	0.12	0.00
Biased	-0.13	0.21	**0.58**	0.24	0.02
Boring	-0.09	-0.52	**0.52**	0.32	0.09
Long	-0.05	0.00	0.38	**0.69**	-0.02
Wordy	-0.10	0.04	0.42	**0.58**	-0.15
Complex	0.15	0.24	0.28	**0.39**	0.18
Technical	0.28	0.25	0.30	**0.34**	-0.13
Coherent	0.45	0.01	-0.16	-0.04	**0.56**

*Factor 5 had an eigenvalue of .97, which is under 1 but significantly higher than the eigenvalue of the next highest factor. Factor 5 also only loaded onto a single item (Coherent)

In Table 3, 18 of the 26 variables had a clearly high loading on one of the five factors (i.e. they possessed relatively small loadings on the remaining factors of at least .20 lesser than the highest loading in absolute value) whereas the other eight measures had moderate loadings on more than 1 factor.

Factor Labels

- Factor 1 had high loadings of 14 variables, the highest of any of the factors by a wide margin. Variables such as Honest, Factual, and Fair, are directly tied to the credibility of the source and the message. Additionally, variables like Precise, Thorough, Informative, Conclusive, Concise, Detailed, Moral, Educational, and Timely point toward the overall quality and representativeness of content, encompassing both the depth and clarity of reporting as well as adherence to journalistic standards. As such, this factor was labeled "Quality," reflecting aspects related to the quality of reporting/writing, perceived credibility of the source/message, and representativeness of journalistic norms.
- Factor 2 had high loadings of variables that evoke positive sentiment and a higher level of reader engagement. This factor included Interesting, Thought-provoking, and Amazing, all of which suggest content that is not only attention-grabbing but also stimulates deeper thinking and positive reactions from readers. As such, this factor was labeled "Engagement".
- Factor 3 had high loadings of variables that suggest content might be off-putting, strange, or difficult to relate to. Variables like Weird, Rote, Biased, and Boring indicate content that either feels unfamiliar, one-sided, or lacking in engagement, potentially leading to a sense of alienation or disconnection for the reader. Thus, the factor was aptly labeled "Alienation".
- Factor 4 (labeled Effort) had high loadings of variables such as Long, Wordy, Complex, and Technical, which point towards content that requires more effort to read and

comprehend. These variables suggest that the articles might be verbose or complex, necessitating a higher level of effort from readers to parse through the information.

- Factor 5 had a single high loading from the variable Coherent, indicating the logical flow and clarity of the content. This factor highlights the importance of coherence in how readers perceive and evaluate AI-generated news, with a focus on the logical structure and understandability of the content.

The analysis of secondary factor loadings, in addition to primary high loadings, offers deeper insights into the nuanced relationships between the higher level factors underlying the data. Among the variables that primarily loaded onto Quality, several also showed moderate loadings on Engagement. This suggests a relationship where content deemed high in quality (accurate, fair, believable) also engages readers. Specifically, attributes like being Precise, Conclusive, Concise, Thorough, Informative, Detailed, Educational, and Timely might not only contribute to the perceived quality of the content but also enhance its engaging nature.

Thorough, Informative, and Detailed variables having moderate loadings on Effort indicates that while these attributes enhance quality and engagement, they also require more cognitive effort from the readers. The depth and detail of the content might demand higher concentration and processing, thus impacting how effortless or taxing the reading experience is. The moderate loading of Moral on Alienation could be explained by the perception that discussions of morality in news content, especially when generated by AI, might create a sense of discomfort or ethical ambiguity.

In the Engagement category, Amazing having a moderate loading on Alienation suggests that while the content is engaging and perhaps surprising, it might also be perceived as overly sensational or unrealistic when attributed to AI, leading to a sense of alienation or skepticism. For variables under Alienation, the fact that Boring has an equally high negative loading on Engagement is quite telling. It underscores an inverse relationship where content perceived as boring not only fails to engage but actively disengages or alienates the audience. Additionally, its moderate loading on Effort implies that boring content might also be seen as requiring unnecessary or unfruitful effort to engage with.

All variables under Effort having moderate loadings on Alienation suggests a relationship where content that is long, wordy, complex, or technical could potentially alienate readers. This might be due to the increased effort required to understand such content, which could lead to frustration or disengagement, especially if readers do not immediately see the value or relevance of investing their time and cognitive resources. Finally, Coherent having a moderate loading on Quality reinforces the idea that clarity and logical structure are not only crucial for understanding (coherence) but also contribute significantly to the perceived overall quality of the content.

Table 4

Factor Loadings of News Stories Ratings - Oblimin Rotation

	Factor 1: Trust	**Factor 2: Prolixity**	**Factor 3: Engagement**	**Factor 4: Substance**	**Factor 5: Clarity**	**Factor 6: Alienation**	**Factor 7: Complexity***	**Factor 8: Effort**	**Factor 9: Neutrality***

Honest	**0.81**	-0.01	0.05	0.03	0.03	0.02	0.00	0.06	0.08
Factual	**0.71**	0.00	-0.03	0.11	0.09	0.00	0.01	-0.02	-0.06
Believable	**0.63**	-0.09	-0.02	0.11	-0.01	0.00	0.05	-0.18	0.00
Fair	**0.54**	-0.04	0.08	-0.08	0.01	-0.10	0.10	-0.08	0.32
Moral	**0.42**	0.06	0.05	-0.04	0.14	0.28	-0.05	0.05	0.15
Long	-0.04	**0.86**	-0.05	-0.05	-0.01	-0.04	0.07	-0.05	0.02
Wordy	0.00	**0.70**	-0.05	0.06	-0.10	0.11	-0.05	0.11	-0.02
Boring	-0.02	0.17	**-0.73**	0.04	-0.01	0.20	0.10	-0.01	0.04
Interesting	0.04	0.01	**0.61**	0.16	0.08	0.22	0.16	-0.01	0.06
Thought-provoking	-0.04	0.14	**0.39**	0.26	0.10	0.21	0.12	-0.02	-0.04
Amazing	0.10	0.24	**0.33**	0.08	0.14	0.24	0.05	0.26	-0.08
Educational	-0.03	-0.04	0.01	**0.71**	0.03	0.05	0.06	0.03	0.09
Informative	0.24	0.01	0.05	**0.62**	0.00	-0.10	0.01	-0.05	-0.02
Concise	0.02	-0.17	0.04	-0.02	**0.71**	0.11	0.01	-0.08	0.07
Precise	0.22	0.07	0.08	0.15	**0.48**	-0.11	0.00	0.00	0.04
Conclusive	0.14	0.10	0.03	0.25	**0.41**	-0.06	-0.09	0.07	0.02
Thorough	0.23	0.22	0.06	0.20	**0.32**	-0.26	0.09	0.04	0.05
Detailed	0.07	0.26	0.11	0.20	**0.30**	-0.26	0.17	-0.03	0.02
Biased	0.05	0.22	-0.05	-0.02	0.06	**0.53**	0.15	0.02	-0.23
Weird	-0.22	0.17	-0.06	-0.13	0.07	**0.39**	0.11	0.33	0.16
Rote	-0.06	0.01	-0.28	0.13	0.15	**0.38**	0.11	0.17	0.04
Complex	0.00	0.02	0.01	0.00	-0.06	0.01	**0.81**	-0.03	-0.02

Coherent	0.12	-0.01	-0.06	0.03	0.25	0.00	0.09	**-0.56**	0.09
Technical	0.10	0.00	-0.06	0.11	0.14	-0.10	0.41	**0.41**	0.00
Neutral	0.08	0.03	-0.05	0.07	0.02	-0.03	-0.04	0.00	**0.69**
Timely	0.00	0.02	0.18	0.23	0.11	0.07	0.08	-0.25	0.26

*Factor 7 loaded onto one item (Complexity)
**Factor 9 loaded onto one item (Neutrality)

After applying an oblimin rather than a varimax rotation, resulting in the data in Table 4, 16 of the 26 variables had a clearly high loading on one of the nine factors (a difference of at least 0.20 between the highest and second-highest absolute values of each loading from the variable). Conversely, 9 of the 26 factors had a moderately high loading in addition to a clearly high loading. One variable (Timely) failed to saliently load onto any factor.

Factor Labels - Oblimin Rotation

- Factor 1 has high loadings of five variables integral to the trustworthiness and reliability of news content: Honest, Factual, Believable, Fair, and Moral. This factor reflects attributes key to establishing the credibility of news. These variables suggest that readers place high value on authenticity, accuracy, and ethical considerations in news content, especially when generated by AI. This factor is aptly labeled “Trust”.
- Factor 2 has high loadings of variables Long and Wordy. This factor captures the verbosity and lengthiness of content. It reflects reader perceptions of articles that may be overly detailed or extended, potentially affecting the readability and accessibility of the news. This factor is thus named “Prolixity”.
- Factor 3 had high loadings of four variables representing the ability of the content to engage or disengage the reader. This factor highlights the importance of keeping the audience captivated and mentally stimulated, leading to its label as “Engagement” similar to the factor in the varimax results.
- Factor 4, having high loadings of the two variables Educational and Informative, emphasizes the informative value and educational quality of the news content. It reflects a preference for content that is enriching and enlightening, and is therefore labeled “Substance.” The term Substance encapsulates the depth, richness, and informative nature of the content, aligning well with the two variables that load onto the factor.
- Factor 5 possessed high loadings of the variables Concise, Precise, Conclusive, Thorough, and Detailed. This fifth factor pertains to the craftsmanship of the writing, focusing on clarity, accuracy, and depth in news presentation, and leading to its designation as “Clarity.” This label emphasizes the clear, concise, and precise nature of the writing style.
- Factor 6, having high loadings of the variables Biased, Weird, and Rote, reflects elements that might create a sense of estrangement or disconnection for the reader. It is appropriately termed “Alienation,” and is similar in structure to the factor from the varimax analysis.

- Factor 7, having a high loading solely of the variable Complex, captures the intricacy or sophistication of the content. This factor indicates the influence of complexity on understanding and engagement with the news and is labeled “Complexity.”
- Factor 8 had its highest loadings on Coherent and Technical. This factor relates to the mental effort required by readers to comprehend and engage with the content. Emphasizing clarity and technicality in news presentation, this factor is named “Effort,” and is constructed similarly to the factor in the varimax analysis.
- Factor 9, consisting of just Neutral, points to the impartiality and unbiased nature of the news, and is thus labeled “Neutrality.”

As previously mentioned, the analysis of secondary factor loadings, in addition to primary high loadings, offers deeper insights into the nuanced relationships between the higher level factors underlying the data. Among the variables with high Trust loadings, the additional moderate loading of Moral onto Alienation is notable in its consistency with the varimax results. This loading again suggests a complex relationship between ethical considerations and feelings of disconnection. When AI-generated content addresses moral issues, it might raise concerns or skepticism among readers about the AIs ability to navigate complex ethical landscapes, potentially leading to alienation. For the Engagement loadings, Thought-provoking loading moderately onto Substance and Alienation indicates that while engaging content stimulates deeper thinking, it might also touch on complex or sensitive topics that can alienate some readers. Amazing showing moderate loadings on Prolixity, Alienation, and Effort is intriguing. It suggests that while such content is captivating, it might also be perceived as verbose, potentially alienating or requiring more cognitive effort to process. The similar loading onto Alienation in the varimax analysis reinforces this idea.

For the Clarity loadings, Conclusive, Thorough, and Detailed having moderate loadings on Substance align with the notion that clear writing often goes hand-in-hand with substantive content. The moderate loading of Thorough and Detailed on Prolixity and their negative loading on Alienation suggest a balance between depth and accessibility; while thorough and detailed reporting is valued for its substance, there is a risk of it becoming verbose and potentially alienating if not presented clearly. For the Alienation loadings, Weird loading moderately on Effort and negatively on Trust might reflect a perception that unconventional or unusual content, while intriguing, can be challenging to comprehend and might undermine the perceived credibility of the content. Rote showing a negative loading on Engagement suggests that content perceived as mundane or formulaic is not only alienating but also fails to engage readers effectively.

For the Effort loadings, Technical showing an equal loading on Complexity indicates a close relationship between the technical nature of content and its complexity, both contributing to the effort required in understanding the material. Timely loading moderately on Substance but negatively on Effort could suggest that while timely content is valued for its relevance and substance, it might be presented in a way that requires less cognitive effort, perhaps due to its immediacy or the nature of its presentation.

An oblimin rotation assumes factors are not independent and are correlated; as such, the strength of correlations between factors are a worthy subject of analysis. Table 5 shows the values of the factor correlations.

Table 5
Factor Correlations - Oblimin Rotation

	Trust	Prolixity	Engagement	Substance	Clarity	Alienation	Complexity	Effort	Neutrality
Trust	1.00	-0.04	0.30	0.62	0.57	-0.20	0.21	-0.23	0.42
Prolixity	-0.04	1.00	-0.17	0.24	0.11	0.29	0.48	0.34	-0.07
Engagement	0.30	-0.17	1.00	0.38	0.31	-0.08	0.06	-0.05	0.04
Substance	0.62	0.24	0.38	1.00	0.56	-0.03	0.42	-0.01	0.19
Clarity	0.57	0.11	0.31	0.56	1.00	0.11	0.35	-0.03	0.35
Alienation	-0.20	0.29	-0.08	-0.03	0.11	1.00	0.24	0.29	-0.04
Complexity	0.21	0.48	0.06	0.42	0.35	0.24	1.00	0.12	0.10
Effort	-0.23	0.34	-0.05	-0.01	-0.03	0.29	0.12	1.00	-0.14
Neutrality	0.42	-0.07	0.04	0.19	0.35	-0.04	0.10	-0.14	1.00

The two strongest correlations were Factor 1 (Trust)-Factor 4 (Substance) and Factor 1 (Trust)-Factor 5 (Clarity). This is logically consistent with the varimax rotation, given the combination of trust-themed adjectives and quality-themed adjectives were combined in the first factor from that analysis. The next-strongest correlation was between Factor 4 (Substance) and Factor 5 (Clarity) themselves.

The Alienation factor emerged as a particularly salient and inter-correlative element in the analysis, underscoring its significance in shaping reader perceptions of AI-generated news content. A notable trend observed was the common occurrence of variables with high loadings across multiple factors also sharing moderate loadings on the Alienation factor. This pattern suggests that while certain attributes of the news content are positively associated with factors like Trust, Engagement, or Clarity, they simultaneously hold the potential to alienate readers, perhaps due to the unconventional or off-putting nature of the content. The moderate negative loading of certain variables onto Trust in the obliquely rotated analysis further accentuates this point. It indicates the presence of a negative correlation between content generated by AI and the way in which its content contributes to feelings of alienation. These feelings could be potentially due to perceived biases or the impersonal nature in which AI-generated content reads (or primes to read) to consumers.

The finding that the Biased variable loaded highest onto the Alienation factor–rather than a factor typically associated with credibility or quality–in both factor structures (varimax and oblimin) is intriguing. There are two different considerations that might explain this phenomenon; the first being the influence of social AI perceptions on content perceptions, and the second being specific source- and content-based cues that relate to alienation. When readers know that a news article is generated by AI, their perception of bias could be influenced by preconceived notions about artificial intelligence. The loading of Biased onto Alienation might reflect broader societal concerns about the role of AI in media and information dissemination. As AI becomes more prevalent in journalism, there are growing discussions and apprehensions about AIs role in shaping narratives or perpetuating biases. This societal context could influence how readers interpret and react to AI-generated content, particularly regarding bias. There is often a skepticism about AIs ability to be truly neutral or unbiased, as AI systems can inadvertently reflect the biases present in their training data. This skepticism could lead to a heightened sensitivity to any perceived bias in AI-generated articles, which might contribute to a feeling of alienation rather than just being a mark against perceived quality.

The concept of alienation in this context could be tied not just to the content of the news but also to its source. If readers perceive AI-generated content as inherently biased, regardless of its actual neutrality or balance, this perception could lead to a sense of disconnect or mistrust. The "machine" behind the news might be seen as less capable of fair and balanced reporting compared to a human journalist, contributing to a sense of alienation. The label of bias in a news article, especially one attributed to AI, might trigger stronger emotional reactions compared to other quality-related concerns. This reaction could lead to feelings of discomfort, distrust, or disagreement, aligning more closely with the theme of alienation. In this sense, bias isn't just a marker of quality but becomes a barrier to reader engagement with the content.

Conclusion

This research reveals that reader perceptions of AI-generated news content differ somewhat from traditional news factors. The introduction of constructs such as Engagement and Alienation, robust and drawing from elements of legacy constructs such as Quality in addition to unique descriptors provided in this survey, demonstrates the importance of reassessing news perception criteria given the advent of generative AI in journalism. Furthermore, the factor analysis identifies the potential importance of Trust as a more salient term than credibility in thinking about reader perceptions: when presented with content that has ostensibly been generated with AI, readers may rely on thinking with their heart rather than their head, reflecting the distinction between trustworthiness and credibility.

Future research should investigate the perception differences between human and machine-written content across different article categories. Additionally, future research could use the newly discovered factors (e.g. Effort, Substance, Clarity) in reader perception studies comparing human-written and AI-written news. This research provides valuable insights into the evolving landscape of automated journalism and its perception among news readers.

One limitation of this study involves the cumulative variance explained by the factor structures analyzed. While the identified factors provide meaningful insights into how readers perceive AI-generated news content, they do not account for the entire variance within the dataset. A substantial portion of the variance could be attributed to elements beyond these unifying factors, such as reader-specific preferences, education levels, personal interests, intelligence, and prior knowledge. These individual differences play a critical role in shaping how readers interact with and interpret news content, suggesting that the factor analysis captures only a part of the broader picture.

Additionally, the inherently exploratory nature of this study represents both a limitation and an opportunity for future research. While it offers a foundational understanding and opens avenues for exploring reader perceptions of AI-generated news, the exploratory approach means that the findings are preliminary and should be used to inform, rather than conclusively define, subsequent in-depth investigations. Future research could utilize the factors explored in this study as metrics upon which AI- and human-generated news content (and perhaps communication content more broadly) can be compared. Ultimately, this study provides a foundation for offering a more comprehensive understanding of the multifaceted nature of news perception in the age of AI.

References

Associated Press News (2023). Alibaba: Fiscal Q4 Earnings Snapshot. *Associated Press.* https://apnews.com/article/earnings-0f8801833d9ed13d9c3b5273b2138185.

Bandi, Ajay, Pydi Venkata Satya Ramesh Adapa, and Yudu Eswar Vinay Pratap Kumar Kuchi. "The power of generative AI: A review of requirements, models, input–output formats, evaluation metrics, and challenges." *Future Internet* 15, no. 8 (2023): 260.

Beckett, C. (2023, 18 Sep). Preparing for the coming wave of generative AI in journalism. *Impact of Social Sciences Blog*. https://blogs.lse.ac.uk/impactofsocialsciences/2023/09/18/preparing-for-the-coming-wave-of-generative-ai-in-journalism/#:~:text=Is%20generative%20AI%20changing%20the,a%20variety%20f%20content%20outright.

Biswal, S. K., & Kulkarni, A. J. (2024). *Exploring the Intersection of Artificial Intelligence and Journalism: The Emergence of a New Journalistic Paradigm*. Taylor & Francis.

Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning phrase representations using RNN encoder-decoder for statistical machine translation. *arXiv preprint arXiv:1406.1078*.

Danzon-Chambaud, S. (2022). *Automated news in practice: changing the journalistic doxa during COVID-19, at the BBC and across media organisations* (Doctoral dissertation, Dublin City University).

Dergaa, I., Chamari, K., Zmijewski, P., & Saad, H. B. (2023). From human writing to artificial intelligence generated text: examining the prospects and potential threats of ChatGPT in academic writing. *Biology of sport*, *40*(2), 615.

Diakopoulos, N. (2023). What Could ChatGPT Do for News Production. *Generative AI in the Newsroom*.

Espejel, J. L., Ettifouri, E. H., Alassan, M. S. Y., Chouham, E. M., & Dahhane, W. (2023). GPT-3.5, GPT-4, or BARD? Evaluating LLMs reasoning ability in zero-shot setting and performance boosting through prompts. *Natural Language Processing Journal*, *5*, 100032.

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological methods*, *4*(3), 272.

Galtung, J., & Ruge, M. H. (1965). The structure of foreign news: The presentation of the Congo, Cuba and Cyprus crises in four Norwegian newspapers. *Journal of peace research*, *2*(1), 64-90.

Graefe, A., & Bohlken, N. (2020). Automated journalism: A meta-analysis of readers' perceptions of human-written in comparison to automated news. *Media and communication*, *8*(3), 50-59.

Harcup, T., & O'neill, D. (2017). What is news? News values revisited (again). *Journalism studies*, *18*(12), 1470-1488.

Harris, Z. S. (1954). Distributional structure. *Word*, *10*(2-3), 146-162.

Henestrosa, A. L., Greving, H., & Kimmerle, J. (2023). Automated journalism: The effects of AI authorship and evaluative information on the perception of a science journalism article. *Computers in Human Behavior*, *138*, 107445.

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, *9*(8), 1735-1780.

Jordan, M.I. (1985). *The Learning of Representations For Sequential Performance.* University of California, San Diego.

Joye, S., Heinrich, A., & Wöhlert, R. (2016). 50 years of Galtung and Ruge: Reflections on their model of news values and its relevance for the study of journalism and communication today. *CM. Communication and Media*, 11(36), 5-28.

Kunert, J. (2020). Automation in sports reporting: Strategies of data providers, software providers, and media outlets. *Media and Communication*, *8*(3), 5-15.

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*.

Moss, A. J., Rosenzweig, C., Robinson, J., Jaffe, S. N., & Litman, L. (2023). Is it ethical to use

Mechanical Turk for behavioral research? Relevant data from a representative survey of MTurk participants and wages. *Behavior Research Methods*, *55*(8), 4048-4067.

Nunnally, J. C. (1978). An overview of psychological measurement. *Clinical diagnosis of mental disorders: A handbook*, 97-146.

Panagia, D. (2021). On the possibilities of a political theory of algorithms. *Political Theory*, *49*(1), 109-133.

Robinson, J. P., & Levy, M. R. (1986). Interpersonal communication and news comprehension. *Public Opinion Quarterly*, *50*(2), 160-175.

Sundar, S. S. (1999). Exploring receivers' criteria for perception of print and online news. *Journalism & Mass Communication Quarterly*, *76*(2), 373-386.

Sundar, S. S., & Kim, J. (2019, May). Machine heuristic: When we trust computers more than humans with our personal information. In *Proceedings of the 2019 CHI Conference on human factors in computing systems* (pp. 1-9).

Sundar, S. S., Oeldorf-Hirsch, A., & Garga, A. (2008, October). A cognitive-heuristics approach to understanding presence in virtual environments. In *PRESENCE 2008: Proceedings of the 11th Annual International Workshop on Presence* (pp. 219-228). Padova, Italy: CLEUP Cooperativa Libraria Universitaria Padova.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, *30*.

Wang, S., & Huang, G. (2024). The impact of machine authorship on news audience perceptions: a meta-analysis of experimental studies. *Communication Research*, 00936502241229794.

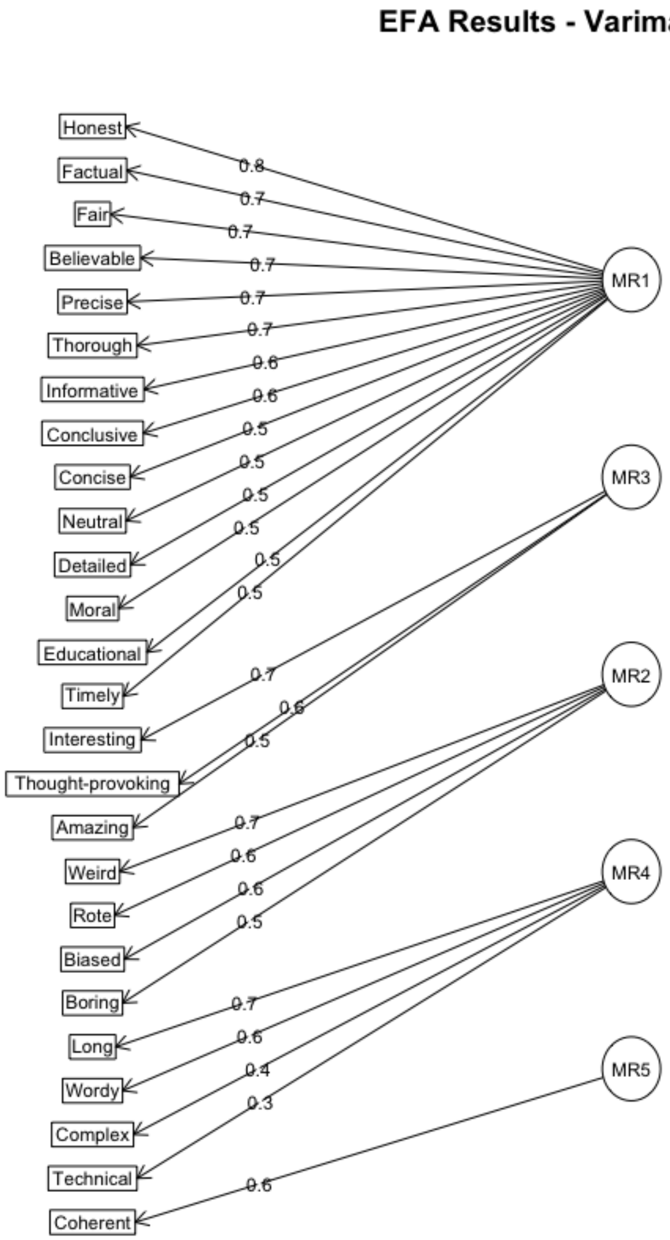
Appendix

Unfiltered List of Descriptors (Phase 1)

1. **Amazing**
2. **Balanced**
3. **Believable**
4. **Boring**
5. **Complete**
6. **Complex**
7. **Concise**
8. Confrontational - appropriateness (reaction to subject matter)
9. Decisive - redundancy (conclusive)
10. Definitive - redundancy (conclusive)
11. **Detailed**
12. Digital - appropriateness (reaction to subject matter - digital rights act article)
13. Digital - appropriateness (digital rights act article)
14. Disappointing - appropriateness (reaction to subject matter)
15. **Educational**
16. Fabulous - redundancy (amazing)
17. **Factual**
18. Factual - duplicate
19. Flowing - **changed to Coherent**
20. Great - redundancy (amazing)
21. Greedy - appropriateness (reaction to subject matter)
22. **Honest**
23. **Inconclusive (Conclusive)**
24. **Informative**
25. Informative - dup
26. Informative - dup
27. Informative - dup
28. Informative - dup
29. **Interesting**
30. Interesting - dup
31. Lengthy - redundancy (long)
32. Limiting - appropriateness (reaction to subject matter)
33. **Long**
34. Long - duplicate
35. Long - duplicate
36. **Moderate**
37. **Moral**
38. **Neutral**
39. New - redundancy (timely)
40. Possible - appropriateness (subj matter)
41. **Precise**
42. **Rote**
43. Sad - appropriateness (subj matter)

44. Technical
45. Thorough
46. Thought-provoking
47. Thought-Provoking - duplicate
48. Threatening - appropriateness (subj matter)
49. Timely
50. Unbelievable - redundancy (believable)
51. Unbiased (Biased)
52. Unbiased - duplicate
53. Uncompromising - appropriateness (subj matter)
54. Unfair - redundancy
55. Uninteresting - redundancy
56. Verbose - redundancy
57. Weird
58. Wordy
59. Wrong - redundancy (factual)

Figure 1: EFA Results Diagram - Varimax



EFA Factors - Varimax
Factor 1 - **Quality**
Honest
Factual
Fair
Believable
Precise
Thorough
Informative
Conclusive
Concise

Neutral
Detailed
Moral
Educational
Timely

Factor 2 - **Engagement**
Interesting
Thought-provoking
Amazing

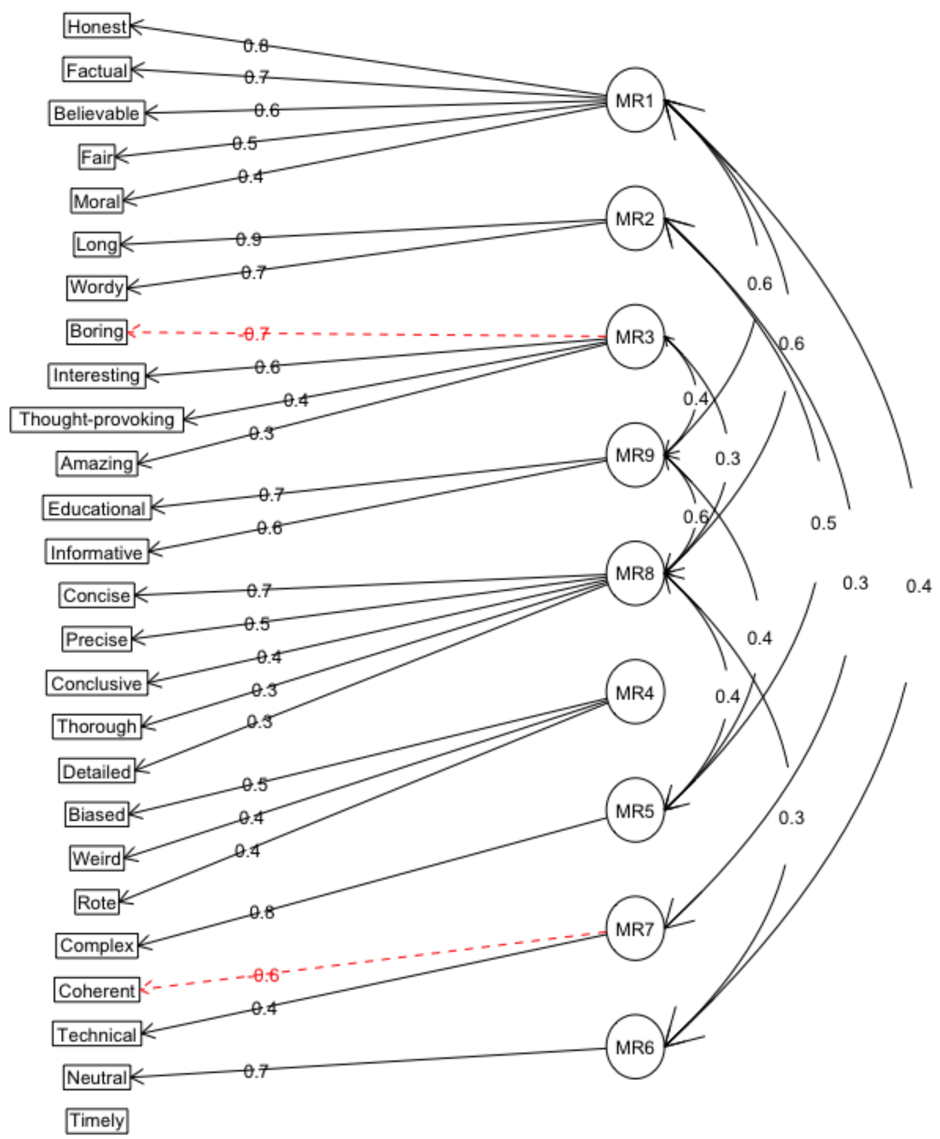
Factor 3 - **Alienation**
Weird
Rote
Biased
Boring

Factor 4 - **Effort**
Long
Wordy
Complex
Technical

Factor 5 - **Coherence**
Coherent

Figure 2: EFA Results Diagram - Oblimin

EFA Results - Oblimin Rotation



EFA Factors - Oblimin

<u>Factor 1 (MR1) - Credibility</u>
Honest
Factual
Believable
Fair
Moral

<u>Factor 2 (MR2) - Prolixity</u>
Long

Wordy

Factor 3 (MR3) - Engagement
Boring
Interesting
Thought-provoking
Amazing

Factor 4 (MR9) - Substance
Educational
Informative

Factor 5 (MR8) - Clarity
Concise
Precise
Conclusive
Thorough
Detailed

Factor 6 (MR4) - Alienation
Biased
Weird
Rote

Factor 7 (MR5) - Complexity
Complex

Factor 8 (MR7) - Effort
Coherent
Technical

Factor 9 (MR6) - Neutrality
Neutral

No factor: Timely