Refining automated news? The role of journalistic post-editing in shaping reader perceptions

Sina Thäsler-Kordonouri [ID]
Ludwig-Maximilians-Universität München (LMU Munich), Germany

Abstract

Automating news production can change the composition of news output and negatively affect its evaluation by the readership. Some journalists decide to post-edit automated news texts to improve their quality for the reader – a practice that, although essential to automated journalism, is still largely underexplored. Using a large-scale online experiment with UK online news consumers ($N = 4734$), this study investigates how readers perceive the changes made during post-editing and whether they achieve the journalistic goal of making the automated articles more appealing. Results show that, regarding the criteria examined, post-editing has no discernible influence on readers' perception of automated articles. However, compared to manually written texts, it appears that this practice may not always be necessary to receive a positive rating from readers. These findings are discussed in the context of the evolving relationship between journalists and technology in news production and local journalism's decreasing resources.

Keywords

Automated journalism, post-editing, local journalism, online experiment, AI journalism

Introduction

As technological capabilities evolve, the number of newsrooms automating their news production is growing rapidly. This practice, called automated journalism, integrates powerful algorithmic systems into the news value chain to increase production efficiency (Haim and Graefe, 2017; Leppänen et al., 2017). Using such systems changes how

Corresponding author:
Sina Thäsler-Kordonouri, Department of Media and Communication, Ludwig-Maximilians-Universität München (LMU Munich), Munich 80538, Germany.
Email: sina.thaesler-kordonouri@ifkw.lmu.de

journalists approach news writing, as they have to adapt their workflow to the systems' logic (Thäsler-Kordonouri et al. 2024). Accordingly, automated journalism is an example of the hybridisation of journalism, in which journalists progressively share agency with technological actants in news production (Dörr and Hollnbuchner, 2017; Lewis and Westlund, 2015; Porlezza and Di Salvo, 2020).

Several studies with readers and journalists show that automating news production can affect the perceived quality of news reporting (Wang and Huang, 2024; Diakopoulos, 2020; Graefe and Bohlken, 2020; Thurman et al., 2017). Accordingly, journalists often downplay these algorithmic systems' skills in news writing, thus showing tendencies of reactance regarding their potential as a professional threat (Beckett and Yaseen, 2023). These attitudes are further substantiated by practices such as the post-editing of automated news texts – that is, the journalistic revision of the narratives of automatically generated texts for the purpose of their qualitative improvement – which help manifest journalistic influence in hybrid news production (Thäsler-Kordonouri, 2024).

Post-editing and its effects on the readership are still largely underexplored. In an earlier study, I identified journalistic post-editing strategies that specifically address the editorial shortcomings journalists believe to be caused by automating the news production workflow (Thäsler-Kordonouri, 2024). The present study investigates which, if any, strategic editing of automated news texts on the part of journalists is recognised by the readership and, if it is, what its effects might be. Using data from a large-scale online experiment with 4734 British online news consumers, the present study analyses (1) how readers perceive the editorial changes made during post-editing and (2) whether these changes achieve the journalistic goal of making the automated articles more appealing, in particular more likeable and easier to understand. To contextualise reader ratings of automated and post-edited articles, their evaluations are compared to those of manually written stories, which have been shown to perform better in several relevant perception categories, including quality, readability (Graefe and Bohlken, 2020), and credibility (Wang and Huang, 2024).

This study advances research on automated journalism by analysing readers' perceptions of journalistic post-editing. The implementation of this editorial practice, which journalists consider important in their engagement with automated news production, is discussed and evaluated. Thus, this study aims to inform journalistic strategies for integrating and handling automated news production (Wilner et al., 2024).

Automation's agency in hybrid news production

Given technological advances, the demand for rapid reporting, and limited resources in newsrooms, news production is increasingly complemented by algorithmic systems that enable journalists to efficiently source, analyse, and process publicly available data for reporting (Borges-Rey, 2020; Tandoc and Oh, 2017). Journalistic work realised with the 'advanced application of computing, algorithms, and automation' (Thurman, 2019: 180) has been described as 'innovative [information] processing that occurs at the intersection between journalism and data technology' (Gynnild, 2014: 715). Although this is not a new phenomenon, the algorithmic systems' rapidly evolving

computing power and capabilities offer increasing opportunities to be integrated into editorial processes (Guzman and Lewis, 2020; Thurman, 2019).

This integration is consequential for journalistic practice, as the workings of the algorithmic systems at various levels – including the data processed for training the model, the setup of the model, the way information is processed through the model, and the way information is presented to the journalist – can affect the content and presentation of reported information in news output (Diakopoulos and Koliska, 2017). Therefore, close monitoring of the programming and configuration of algorithmic systems used in news production is essential to ensure that their design and output are consistent with the ethical principles of journalism (Dörr and Hollnbuchner, 2017).

This 'algorithmic revolution in knowledge production' (Anderson, 2012: 1006) shows how journalism has entered a phase of hybridity, which has been interpreted as an indication of the permeability and low autonomy of some actors in the journalistic field vis-à-vis the non-journalistic organisations responsible for the development of the algorithmic systems used in newsrooms (Danzon-Chambaud and Cornia, 2021; Leppänen et al., 2017; Porlezza, 2023; Simon, 2022). Some researchers argue in favour of the transformative potential of these systems for journalism, as the more they are integrated into news production workflows, the more journalists' 'skills, tasks and jobs will necessarily evolve, most likely to privilege abstract thinking, creativity, and problem-solving' (Diakopoulos, 2019: 131).

The 'Agents of Media Innovations' (AMI) approach conceptualises the reconfiguration of agency in the news value chain that can come along hybrid forms of news production, such as automated journalism (Lewis and Westlund, 2015). This approach applies a socio-technical understanding of news organisations and the dynamics that affect their activities by acknowledging 'the extent to which contemporary journalism is becoming interconnected with technological tools, processes, and ways of thinking as the new organizing logics of media work' (p. 21). The AMI emphasises the relevance of considering non-editorial actors and actants and those external to the news sector, such as business leaders, policymakers, or audiences, when investigating the factors that shape news coverage and other related news media activities. Using cross-media news work as an example, Lewis and Westlund (2015) discuss how the AMI may help 'reveal nuances in the relationships among human actors inside the organization, human audiences beyond it, and the nonhuman actants that cross-mediate their interplay' (p. 21).

To highlight the influence of each leading agent in creating and configuring news media activities, the authors use a matrix logic that reflects the extent to which editorial and non-editorial actors, technology or audiences may play the leading role. 'These activities are routinised practices and patterns of action through which an organisation's institutional logic is made manifest through media' (Kuai et al., 2022: 1896). Alongside other types of agents involved in news production, the AMI emphasises the influence of technological actants on the creation of media output; however, without taking on a deterministic understanding that assumes technology takes over control in news production. Instead, the approach aims to manifest the outcome of 'human–technology tension [that is] best understood as a continuum between manual and computational modes of orientation and output in contemporary cross-media news work – a way of perceiving

the relative gravitational pull of each dimension in shaping news publishing' (Lewis and Westlund, 2015: 20). Thereby, agency associated with technological actants in news production is not of the sort held by human actors with consciousness and free will (see Latour, 2005). Instead, agency is attributed to technology and manifests in 'recognizing the constraints that humans may face in working within technical systems of ever-growing complexity and ubiquity' (Lewis and Westlund, 2015: 21).

Overall, the AMI approach emphasises the increasing attribution of agency to non-editorial agents involved in news production, including technological actants and the audience (or various representations thereof), alongside journalists as traditional agency-holders in editorial processes. In the case of automated news production, taking on the AMI approach suggests that the composition of the news story produced with the help of automation is not only characterised by the actions of human actors (journalists), but also by the functioning and affordances of technological actants (automation systems), and the demands of the audience (readership). The following sections theoretically discuss these agents' interplay and their impact on the composition of news messages in automated journalism.

Automating the news production workflow: Technological actants' potential impact on the composition of news

Automated journalism has been defined as a process in which algorithms are used to 'convert data into narrative news texts' (Carlson, 2015: 417) with varying degrees of human input beyond the initial programming (Thurman et al., 2024; Waddell, 2019). In automated journalism, editorial tasks are realised with the help of natural language generation (NLG) models. These models can have varying degrees of computational complexity, ranging from rule-based variants that use manually generated templates into which data is automatically inserted (Diakopoulos, 2019; Leppänen et al., 2017) to variants that use machine learning (Danzon-Chambaud, 2023). Regardless of the technological complexity, automated journalism represents a fundamental integration of technological actants (Lewis and Westlund, 2015) into news production. This integration consequently impacts journalists' approaches to creating news articles, their narrative structure, and their composition (Thäsler-Kordonouri et al. 2024). Put differently, it manifests the 'distinct role of technology and the inherent tension between human and machine approaches' (Lewis and Westlund, 2015: 20).

Automated journalism usually requires information in a data-based, machine-readable form (Haim and Graefe, 2017), which limits the possibility of automating the production of complex stories, as data has limitations in representing causalities between events (Caswell and Dörr, 2019; Tandoc et al., 2022). Additionally, without further editorial input from journalists, automation limits the ability to include some relevant editorial components of journalistic reporting, such as contextual information, anecdotes, or quotes (Caswell, 2019). Therefore, automated journalism is mainly deployed in producing relatively simple stories, as 'the methods used for encoding news events and stories as data are only suitable for journalism that is […] formulaic' (Caswell and Dörr, 2019: 954).

With the widely used template-based variant of automated journalism, journalists use NLG software to create text templates into which data is automatically embedded. To be useful for large-scale news production, these templates must be adaptable to various story contingencies. Therefore, they are textually designed to be used with many variants of data expression (Thurman et al., 2017). These considerations impact journalists' approaches when creating story templates and ultimately influence the composition of the resulting news output (Thäsler-Kordonouri et al., 2024). As Diakopoulos (2019) puts it, journalists creating news templates 'need to approach a story with an understanding of what the available data *could* say – in essence, to imagine how the data could give rise to different angles and stories and delineate the logic that would drive those variations' (p. 131). Therefore, journalists involved in originating auto-mated news are required to have skills 'that are very different from traditional news-room skills, including comfort with abstraction and a willingness to work "computationally" instead of solely with writing' (Caswell and Dörr, 2019: 954). Thus, automating news production causes journalistic workflows to adapt to the logic and capabilities of the algorithmic systems used. That is, the workings of NLG software affect how journalists approach the creation of templates and, therefore, news texts.

Studies have examined how the integration of algorithmic systems in the news cycle impacts journalistic work in terms of workflow (Milosavljević and Vobič, 2019; Thäsler-Kordonouri and Barling 2023; Thurman et al., 2017), normative (Hermida and Young, 2017), ethical (Dörr and Hollnbuchner, 2017; Jamil, 2023), and legal aspects (Kuai et al., 2022). In general, scholars emphasise the effects of these systems on jour-nalistic practice when used (Jamil, 2023; Porlezza and Di Salvo, 2020), arguing that the algorithmic intervention that takes place in processes such as automated journalism stands for a 'structural transformation of making news and engaging with the audience' (Helberger et al., 2022: 1606).

Emphasising the audience: Journalistic post-editing of automated news texts

Given that algorithmic systems affect journalistic workflows in automated journalism, the question arises as to whether, and if so, how journalistic work processes in conjunction with the automation systems can produce news output optimally adapted to the reader-ship's needs. To investigate reader perceptions of automated news texts, several experi-mental studies have used manually written stories as a point of comparison. Overall, findings show that in the eyes of the readership (Graefe and Bohlken, 2020) and journal-ists (e.g. Diakopoulos, 2020; Møller et al., 2025), automated news output is often outper-formed by manually written reporting. Several meta-analyses of perception studies found that readers rated automated news articles as less readable, of lower quality (Graefe and Bohlken, 2020), and less credible than manually written ones (Wang and Huang, 2024). However, these differences in rating were often rather small.

Similarly, research has shown that while journalists recognise the potential benefits of automation – for example, in terms of increasing efficiency and freeing up resources – there is still a prevailing view that automation cannot replace human creativity and instinct in news production (Milosavljević and Vobič, 2019; Simon, 2023). This

conviction is also reflected in journalists' assessments of automated news texts, which they have rated as narratively insufficient to fulfil readers' needs without significant human intervention (Diakopoulos, 2020; Thurman et al., 2017). In other words, research suggests that the affordances of NLG systems may impose editorial constraints on news production that can affect the composition of the resulting news output in certain regards from the perspective of both journalists and readers.

Consequently, journalists have stated that additional editorial input can be needed to compensate for the narrative shortcomings of automated news (Thäsler-Kordonouri and Barling 2023). This process, called post-editing, includes all aspects of journalistic editing of automatically generated news text before publication (Thäsler-Kordonouri 2024). Post-editing has similarities to journalistic work performed on other third-party news copy, such as press releases or news agency reports, in the sense that journalists use these news artefacts as a starting point for their reporting and then adapt the narrative to the editorial style of their newsroom and/or their perceptions of the readership's needs (e.g. Boumans et al., 2018). Beyond this, however, this editorial practice addresses the specific narrative characteristics that arise from scaling the news production workflow in an automated way (Caswell, 2019; Diakopoulos, 2019; Thäsler-Kordonouri et al., 2024). Therefore, post-editing can be seen as a practice in which journalists 'are negotiating issues of authority, identity, and expertise in connection […] with the machine-led processes assuming more responsibility for functions traditionally associated with professional control' (Lewis and Westlund, 2015: 22).

How post-editing is carried out depends on the available resources of the newsroom and the news value of a story. In post-editing, journalists have stated that they make minor changes to the automatically generated output or use the output as a 'template' or 'starting point' (Thäsler-Kordonouri and Barling 2023, 12) to develop it into a longer news piece. Using interviews with post-editing journalists and a content analysis, in a previous study I analysed how this process is implemented in local journalism workflows. The results show that post-editing journalists say they reduce the quantity of numbers in the story to avoid data overload for readers, increase the local focus of the narrative to tailor it to the area they report for and add contextual information to explain relevant causalities behind the data and the events it represents. These editorial amendments should help increase the appeal of the automatically generated news text for the readership, especially in locally focused news outlets (Thäsler-Kordonouri 2024).

Research questions

Research suggests that due to the advancing capabilities of algorithmic systems used in automated journalism, greater attention should be paid to the relevance of these systems to news production and output (e.g. Dörr, 2019; Guzman and Lewis, 2020). In automated journalism workflows, editorial practices adapt to the functioning of algorithmic NLG systems, with these adaptations affecting the composition of the resulting news texts (Caswell, 2019; Diakopoulos, 2019; Thäsler-Kordonouri et al. 2024). Thus, research points to technological actants – alongside journalists – taking on agency in automated journalism (Dörr and Hollnbuchner, 2017; Lewis and Westlund, 2015). As a result,

journalists may decide to post-edit automated news texts; a practice that can help manifest journalistic control in automated journalism aimed at compensating the perceived narrative shortcomings resulting from automating news production and improving the articles' appeal to the readership (Thäsler-Kordonouri, 2024).

Based on the theoretical considerations and empirical findings outlined above, this study examines how the influences of journalists and automation systems on news production are reflected in readers' evaluations of the resulting news texts. To this end, this study analyses how readers perceive news texts produced either using template-based automation, using template-based automation and post-edited by a journalist or written manually. The first two research questions focus on the editorial criteria addressed by post-editing as well as the intended outcome of their implementation:

RQ1. How do readers evaluate the editorial criteria addressed by post-editing in automated, post-edited, and manually written news texts?

RQ2. How do readers evaluate the comprehensibility and likeability of automated, post-edited, and manually written news texts?

To determine the relevance of post-editing and derive recommendations for journalists, the study furthermore asks:

RQ3. How do readers' evaluations of the editorial criteria addressed by post-editing affect their liking and comprehensibility perceptions of the news texts?

Methodology

The data used for the analysis in this study derives from a large-scale between-subjects online survey experiment that used a sample representative of UK online news consumers by age and gender ($N = 4734$) (see Thäsler-Kordonouri et al. 2024). The Institutional Review Board of the Social Science Faculty at LMU Munich approved the study procedure for compliance with ethical guidelines. The perception criteria used in the survey were developed in a qualitative pre-study (Stalph et al. 2023). In the survey, participants read published news texts that were produced using template-based automation ($n = 1599$), using template-based automation and post-editing by a journalist ($n = 1593$), or written manually by a journalist ($n = 1542$). The news texts were organised into 12 sets, with every set including three news texts created using each production type, about the same story, based on the same data, and published in the same geographic region at a similar time. They covered a range of 12 topics, including energy, transport, crime, politics, and health (see Table A in Supplemental Material).

Stimuli. The news texts were presented to participants in basic HTML formatting, with authorship (byline) and publisher indications excluded. Participants were only informed at the end of the experiment that some of the articles had been produced using automation. The contents of the texts were not modified before exposure to the participants. The automated news texts came from the news wire service Reporters And Data And Robots (RADAR), which generates locally specific data-driven reporting using template-based automation. The post-edited and manually written texts were found via extensive online research using search queries and text-matching commands. The post-edited articles were only classified as such if UK news outlets had published

them, they were based on an automated RADAR article – that is, they used the same data, were published at a similar time, and had matching text – and if editorial changes had been made to the body of the text, not just the headline. Manually written news texts were only matched with the automated and post-edited ones if they had been published at a similar time by UK news outlets, were about the same story, used the same data, and did not include any automation in their production – which was confirmed in personal correspondence with the articles' authors. Thus, articles were only grouped as a set of three (automated, post-edited, manually written) if they were based on the same data, released around the same time, had the same geographical angle, and were not written in different journalistic styles (e.g. to avoid matching opinion pieces with reports).

Measures. The measures used in the survey experiment were derived from a group interview pre-study with 31 UK news consumers from demographically diverse backgrounds (Stalph et al. 2023). Respondents were shown different selections of 21 sets of English-language, data-driven news articles (56 articles in total), produced with and without the help of automation, covering a range of topics and published by UK and international news outlets. Based on these group interviews, 28 perception criteria were extracted, some of which were the basis for the measures used in this study. The variables included in the analyses thus stem from a larger body of perception criteria investigated, which is why the scale levels used to measure respondents' evaluations differ slightly from each other in some cases. Therefore, scale levels were, at times, re-coded for analysis purposes. This process is labelled at the appropriate points.

Quantity of numbers. This criterion measures readers' perception of all numbers featured in a news text, including exact numbers, percentages, rounded numbers, and analogies, by asking: *What do you think about the volume of the different types of numbers used in the article you read?* Respondents indicated their satisfaction with the quantity of numbers in the reporting on a seven-point scale ranging from −3 (Should have used far fewer), through 0 (Contains about the right amount), to +3 (Should have used far more) ($M = -.77$, $SD = 1.03$).

Level of local focus. This criterion measures readers' perception of the overall local focus of the news text's narrative by asking: *Thinking about the area or areas the article focused on, what did you think about the level of local focus?* Respondents indicated their satisfaction with the level of local focus on a three-point scale ranging from 1 (There is not enough), through 2 (Contains about the right amount), to 3 (There is too much) ($M = 1.85$, $SD = .45$).

Level of national focus. This criterion measures readers' perception of the overall national focus of the news text's narrative by asking: *Thinking about the area or areas the article focused on, what did you think about the level of national focus?* Respondents indicated their satisfaction with the level of national focus on a three-point scale ranging from 1 (There is not enough), through 2 (Contains about the right amount), to 3 (There is too much) ($M = 1.76$, $SD = .52$).

Contextual information. These criteria measure readers' perceptions of some relevant contextual information provided in the news texts. The decision to specifically investigate the perceived *amounts of analysis*, *human angle*, and *interventionism* as part of the contextual information was made for two reasons: (1) across all group interviews in the qualitative pre-study, the respondents mentioned that these criteria were relevant to their understanding of the bigger picture of the reported events (Stalph et al. 2023), and (2) in the previous study's interviews, post-editing journalists repeatedly mentioned some of these criteria as relevant to their goal of providing readers with further contextual information (Thäsler-Kordonouri, 2024).

Analysis. This criterion measures readers' perception of the amount of explanation in the news reporting by asking: Do you think the article should have included more or less of the following, or does it contain the right amount? Analysis, for example, by explaining any possible consequences. Respondents answered on a seven-point scale ranging from −3 (Should have used far less), through 0 (Contains about the right amount), to +3 (Should have used far more) ($M = -1.47$, SD = 1.31).

Human angle. This criterion measures readers' perception of the narrative space given to people involved in the reported events by asking: Do you think the article should have included more or less of the following, or does it contain the right amount? Information about the experiences of the people directly affected. Respondents answered on a seven-point scale ranging from −3 (Should have used far less), through 0 (Contains about the right amount), to +3 (Should have used far more) ($M = -1.57$, SD = 1.28).

Interventionism. This criterion measures readers' perception of the amount of solutions-orientated narrative in the news reporting by asking: Do you think the article should have included more or less of the following, or does it contain the right amount? Discussion of solutions. Respondents answered on a seven-point scale ranging from −3 (Should have used far less), through 0 (Contains about the right amount), to +3 (Should have used far more) ($M = -1.69$, $SD = 1.24$).

Liking. This criterion measures readers' liking of a news story by asking: *How much do you like or dislike the article?* Respondents indicated their overall liking of the news article (see Sundar 1999) on a continuous scale from 0 (I disliked it very much) to 60 (I liked it very much) ($M = 33.29$, $SD = 12.36$).

Comprehensibility. This criterion measures readers' perceptions of the comprehensibility of a news story by asking: *In general, how easy or difficult is it to understand the article overall?* Respondents indicated the perceived overall comprehensibility of the news article (see Sundar 1999) on a continuous scale from 0 (It was very difficult to understand) to 60 (It was very easy to understand) ($M = 44.27$, $SD = 13.45$).

Instrument pre-test. Developmental expert reviewing and cognitive interviews with respondents ($N = 10$) were used to pre-test the questionnaire in October 2022, with this

pre-testing including think-aloud and verbal probing procedures (Willis, 2016). The respondents were presented with a survey prototype that used a total of nine stimuli, of which each respondent read one. Stimuli and respondents were matched according to residential area. Story topics of the stimuli included energy, transport, and health. Afterwards, a soft launch was conducted with approximately 100 respondents to test the survey's technical functionality and measures.

Survey administration and data collection. The survey was fielded by YouGov between 26 January and 1 March 2023 using their proprietary online panel. Participants were pre-screened by the panel provider to ensure they were at least 18 years of age and regularly consumed online news. Each experimental treatment group comprised at least 100 participants, with common quotas set on age and gender for the recruitment of each group. All respondents who entered the final sample passed an attention check (Ruble, 2017).

Sample. The online experiment used a representative sample of monthly UK online news consumers aged 18 or over. Participants were divided into 24 locality-specific treatment groups, with at least 100 participants per group. Participants were only exposed to news reporting that matched their place of residence. Participants had a mean age of $M = 50.66$ years ($SD = 15.76$), with 55% of them identifying as female ($n = 2602$) and 45% as male ($n = 2132$). They mainly worked in full-time positions (45.3%), were pensioners (26%), or worked part-time (12.9%).

Results

The first research question asks how readers evaluate the implementation of the editorial criteria addressed by post-editing in the three news text types (automated, post-edited, and manually written). As some of the variances in the metrically scaled variables were heterogeneous, non-parametric Welch tests and Games-Howell post hoc analyses were conducted to compare the mean scores.

The results show that readers would have preferred slightly fewer numbers and less interventionism in the reporting. Nevertheless, they were slightly, but significantly, more satisfied with the presentation of both narrative elements in the manually written stories than in those produced using automation, including the post-edited ones. Thus, post-editing did not significantly improve readers' perceptions of the quantity of numbers and the amount of interventionism in the automated stories. Regarding the amount of analysis and human angle in the reporting, there is no significant difference in readers' perception between automated, post-edited and manually written stories. Still, readers would have liked to see a little less of these narrative elements in all three types of articles Table 1.
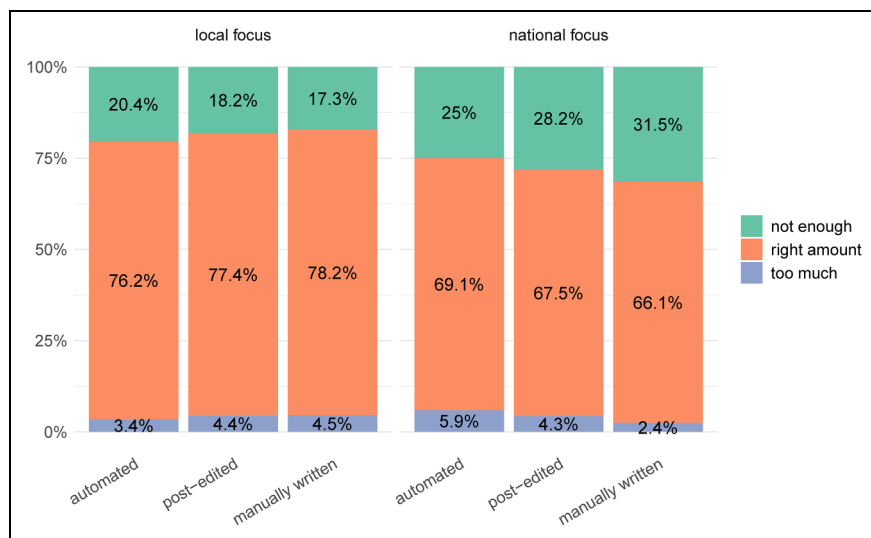
When comparing the categorically scaled variables, results show that there is no significant difference between reader evaluations of the level of local focus between the automated, post-edited, and manually written stories ($\chi^2(4) = 7.23$, $p = .124$). Readers would have wanted a stronger local focus across the three news production types (see Figure 1). However, readers were significantly more satisfied with the amount of national focus in the automated stories compared with post-edited and manually written ones

Table 1. Comparison of reader evaluations of the editorial criteria relevant to post-editing using non-parametric Welch tests (original scales).

Group	Quantity of numbers		Interventionism		Analysis		Human angle	
	M	SD	M	SD	M	SD	M	SD
Automated[a]	−.81	1.02	−1.72	1.25	−1.49	1.33	−1.58	1.27
Post-edited[b]	−.80	1.05	−1.74	1.23	−1.47	1.31	−1.62	1.28
Manually written[c]	−.69	1.02	−1.61	1.25	−1.47	1.31	−1.52	1.27
	$F(2, 3152.53) = 6.12^{**}$		$F(2, 5.16) = 5.16^{**}$		$F(2, 3153.04) = 0.12$		$F(2, 3152.09) = 2.12$	
	$\eta^2 = .003$		$\eta^2 = .002$					
Pairwise comparisons	ab $F(1) = -.01$		ab $F(1) = .03$					
	bc $F(1) = -.11^{**}$		bc $F(1) = .13^{**}$					
	ac $F(1) = -.11^{**}$		ac $F(1) = -.11^{*}$					

Note: *** $p < .001$; **$p < .01$; *$p < .05$.
a,b,c refers to the different groups (automated, post-edited, manually written).

Figure 1. Reader evaluations of the level of local and national focus in the automated, post-edited, and manually written stories.

$(\chi^2(1) = 36.03, \ p < .001, \ \varphi = 0.09.)$ (see Figure 1). Here, the automated production approach slightly but significantly outscored the ones with more human involvement (post-edited, manually written).

To answer RQ2, again, non-parametric Welch tests and Games-Howell post hoc analyses were conducted to compare the mean scores of liking and comprehensibility across the three article-type groups. Results show that there is no significant difference in readers' liking of the automated, post-edited, and manually written stories, which were all liked moderately (see Table 2). However, readers found the manually written stories significantly more comprehensible than those produced with automation, including the post-edited ones. Thus, post-editing did not significantly increase reader perceptions of the comprehensibility of the automated stories.

To answer RQ3, two multiple regression analyses were performed with the editorial criteria addressed by post-editing, regressing on liking and comprehensibility (see Table 3). As variable scale levels deviated slightly and some of the variables had non-linear scales, all variables were recoded to a 0 (negative evaluation/not at all satisfied) to 100 (positive evaluation/completely satisfied) quasi-metric scale according to the procedure recommended by Preston and Colman (2000) before being included in the regression model.

The sampling procedure used in the survey experiment ensured that each experimental group had representative participant distributions in terms of age and gender (see Methodology). However, to control for potential confounding effects caused by personal predispositions of the respondents relevant to their evaluations of the data-driven news reporting, the participants' personal preference for numbers-driven news (statistical affinity) was controlled for.[1]

Table 2. Comparison of readers' liking and comprehensibility perceptions of automated, post-edited, and manually written news texts using non-parametric Welch tests (original scales).

Group	Liking		Comprehensibility	
	M	SD	M	SD
Automated[a]	32.76	12.08	42.92	13.60
Post-edited[b]	33.41	12.07	43.86	13.63
Manually written[c]	33.73	12.93	46.07	12.91
	$F(2, 3144.24) = 2.48$		$F(2, 3153.97) = 23.56$***	
			$\eta^2 = .013$	
Pairwise comparisons			$ab_{comprehensibility} \ F(1) = -.94$	
			$bc_{comprehensibility} \ F(1) = -2.21$***	
			$ac_{comprehensibility} \ F(1) = -3.15$***	

Note: Items used response scales ranging from 0 (I disliked it very much/Very difficult to understand) to 60 (I liked it very much/Very easy to understand).
*** $p < .001$; ** $p < .01$; * $p < .05$.
a,b,c refers to the different groups (automated, post-edited, manually written).

Table 3. Multiple regression analyses of the relationship between the editorial criteria addressed by post-editing and the liking and perceived comprehensibility of the news texts (recoded scales).

Editorial criteria	Liking				Comprehensibility			
	B	β	SE	t	B	β	SE	t
Intercept	36.20***		1.05	34.64	53.81***		1.13	47.81
Quantity of numbers	.12***	.19	.01	13.92	.20***	.30	.01	21.38
Level of local focus	.03***	.06	.01	4.35	.01	.02	.01	1.05
Level of national focus	.01	.02	.01	1.27	.01	.02	.01	1.02
Analysis	.01	.03	.01	1.87	.02**	.04	.01	.00
Human angle	−.00	−.01	.01	−.57	−.00	−.00	.01	−.18
Interventionism	.00	−.00	.01	−.05	.00	.01	.01	.44
Statistical affinity	.16***	.19	.01	13.56	.07***	.08	.01	5.57
R^2	.09***				.10***			
corr. R^2	.09				.10			
F(df)	63.48(7)				78.36(7)			
Durbin–Watson	1.93				1.92			

Note: All variables in this analysis were recoded to scales from 0 (negative evaluation) to 100 (positive evaluation) following Preston and Colman (2000).
*** $p < .001$; ** $p < .01$; * $p < .05$.

Overall, the editorial criteria addressed by post-editing explain about 10% of the variance of both liking and overall comprehensibility when controlling for statistical affinity. With these results, we find ourselves in an area of rather low explanatory power in relation to liking and comprehensibility. This outcome can be attributed to the study's exploratory nature and the unique narrative criteria being scrutinised.

Results show that, for this type of data-driven news reporting, reader satisfaction with the quantity of numbers presented in the news reporting is most crucial to their liking and comprehensibility perceptions of the reporting. Additionally, satisfaction with the level of local focus significantly contributes to liking, whereby satisfaction with the analysis significantly contributes to the reporting's perceived comprehensibility. In contrast, the readers' perceptions of the level of national focus, human angle and interventionism in the narrative do not play a significant role.

Discussion and conclusion

Automating news production is accompanied by an adaptation of human workflows to algorithmic logic (Dörr and Hollnbuchner, 2017). For news production processes, this adaptation can manifest in various ways; for example, journalists who programme templates no longer write single stories but instead have to compose narratives abstractly to anticipate eventualities in data sets (Diakopoulos, 2019). Moreover, this mode of news production may limit the representation of complex causalities in the resulting story (Caswell and Dörr, 2019) and the integration of specific editorial components of journalistic reporting that cannot be presented in a data-based form into its narrative (Caswell, 2019). This can impact the composition of the story and, in some instances, may negatively affect its evaluation by the readership compared to non-automated news (Wang and Huang, 2024; Graefe and Bohlken, 2020). The influence that algorithmic systems can have on the creative journalistic production process in these and other ways, thus, encourages us to rethink the agency constellations that exist between journalistic actors and technological actants in automated news production (Dörr and Hollnbuchner, 2017; Lewis and Westlund, 2015).

In this context, journalists' post-editing of automated news represents a form of actor-led agency negotiation with technological actants in hybrid news production. In other words, post-editing can be seen as some journalists' reaction to the reconfiguration of agency in automated news production and thus to the fact that 'contemporary journalism is becoming interconnected with technological tools, processes, and ways of thinking' (Lewis and Westlund, 2015: 21). Post-editing should not necessarily be interpreted as an act of professional protectionism vis-à-vis these new technologies on the part of journalists. Instead, through post-editing, hybrid news production is reconciled with institutionally driven, established journalistic practices.

Against the backdrop of these reconfigurations of agency in automated journalism, this study examined reader perceptions of post-editing. The study investigated how readers evaluate the specific editorial criteria relevant to post-editing in automated, post-edited, and manually written news texts and how reader satisfaction with these criteria drives their liking and comprehensibility perceptions of the news texts. Thus, this study evaluates the audience's recognition of the post-editing practice and discusses the implementation of the editorial steps journalists deem relevant.

The results show that readers did not notice an improvement in the implementation of the editorial criteria queried after the editing. In other words, post-editing had no influence on readers' evaluation of the articles compared to their automated precursors. One explanation for this could be that the editing done by journalists may not have been effective enough to be noticed by the readership, for example, because what was

considered sufficient by journalists – in terms of the number or extent of changes made to the text – may not have been substantial enough to make a difference in how readers perceived the news texts. This may be related to the strained resources in local journalism, as an ever-decreasing number of journalists have to shoulder the same editorial workload (Costera Meijer, 2020). In UK local journalism, market concentration and news organisations' advertising-driven profit strategies have 'reduced the number of journalists, kept their wages low and impacted on the newsgathering and reporting practices in ways that diminish the range and quality of editorial' (Franklin, 2007: 12).

Alternatively, the editorial changes journalists deem to be relevant for the post-editing of automated news texts (Thäsler-Kordonouri 2024) may not necessarily address the composition criteria that readers pay attention to when consuming these sorts of news texts. Thus, the editorial changes made during post-editing queried in this study (which are based on journalists' assessments of what is relevant when post-editing) may have been 'overlooked' by the readership. Prior research has demonstrated that there can be discrepancies between the readership presumed by journalists and the actual one, which can result in journalists developing an incongruent understanding of the readership's demands and preferences with regard to the composition of news content. As Robinson (2019) put it, 'a central irony of the newsroom is that while many journalists' decisions are made with readers in mind, the audiences for their work often remain unfocused abstractions in their imagination, built on long-held assumptions, newsroom folklore, and imperfect inference' (p. 4). These primary explanations should be further investigated in future research, which should also consider the increasing use of machine-learning based generative AI for news production in addition to rule-based automation systems and its effects on post-editing.

However, when comparing readers' ratings of the editorial criteria relevant for post-editing in texts created with automation, including those post-edited, to those of texts created without automation, results show that post-editing may not always have been necessary. With regard to the amount of analysis, the human angle, the level of local focus in the news texts, and readers' overall liking of these texts, those created with automation received ratings that were not significantly different from those given to manually written texts. Regarding the perceived level of national focus, the automated texts were even slightly better rated than the manually written ones. These results indicate that the automated news production approach effectively meets readers' expectations of this type of local news reporting regarding these editorial features. Thus, post-editing by journalists should be discussed not only in terms of how it is done but also in terms of *when it is necessary*. In view of scarce journalistic resources in local news (Franklin, 2007), these results provide useful insights for journalists who produce this type of news text, including that which is automated, as they shed light on readers' demands of its composition.

Overall, the texts used in this experiment can be characterised as simple local data-driven news reports. This type of news text lends itself to being automated, as it usually conveys information without containing many complex narrative elements (Caswell and Dörr, 2019; Diakopoulos, 2019). Readers' responses suggest that they also perceived these texts as such since they would have preferred fewer numbers and contextual information and a slightly stronger local focus in the automated, post-edited and manually written texts.

Still, the manually written stories slightly but significantly outperformed those produced with automation, including the post-edited ones, regarding the quantity of

numbers, the amount of interventionism featured, and their overall comprehensibility. These results echo previous research, which found that readers often prefer the narrative composition of manually written news texts over that of automatically generated copy (Wang and Huang, 2024; Graefe and Bohlken, 2020), and emphasise the continued relevance and refinement of human involvement in hybrid news production.

In particular, readers' evaluation of the quantity of numbers in the news reporting plays an important role because results show that it drives their liking and comprehensibility perceptions of the texts. Other studies suggest that using numbers in news texts can benefit their reception if done with reason, which also applies to their quantity (Koetsenruijter, 2011). Furthermore, this study found that reader satisfaction with the level of local focus contributes to their liking and satisfaction with the amount of analysis contributes to their perceived comprehensibility of the news texts. Thus, journalists should pay particular attention to these composition features when producing and editing local data-driven news, including that which is automated.

Overall, the results suggest that, from the readers' point of view, the practice of post-editing as a media activity is mainly determined by journalistic actors – including their institutionally shaped professional perceptions – and their engagement with technological actants, that is, the adaptation of their workflow to the logics of the automation systems. The audience, as another relevant type of actor (Lewis and Westlund, 2015), rather seems to be left on the sidelines when it comes to influencing how journalists carry out post-editing as a media activity. This becomes evident, on the one hand, in the fact that the readership does not recognise any discernible differences between the automated and the post-edited texts with regard to relevant evaluation criteria – which may be due to differences in journalists' and readers' perceptions of what degree of editorial change is sufficient to be recognised, or to journalists' incongruent conceptualisation of what editorial criteria readers pay attention to when consuming this type of news (see Robinson, 2019). Additionally, it shows that in some respects, there seems to be a certain misunderstanding on the part of journalists regarding the necessity of post-editing since the automated texts are already sufficiently composed for the readership in various editorial regards, as they are not differently evaluated from manually written articles. These results are particularly unfortunate given that journalists have stated that their motivation for post-editing is to improve the narrative of the automated texts for their readership (Thäsler-Kordonouri 2024).

This study's results thus emphasise the continued relevance and refinement of human involvement in hybrid news production, thereby, advocating readers' relevance as a point of reference in the journalistic engagement with automatically generated news, including in their post-editing. Not least to support the sustainable use of journalistic resources in local newsrooms that use automated journalism. In an increasingly automated news media landscape, a stronger focus on the needs of the readership and their targeted integration into the composition of news may make the decisive difference.

Limitations

It should be noted that this study has its limitations. The research design focused on the investigation of template-based automated journalism. Although this is a widespread method of automating news production, the decision to include only this type of news

automation limits the generalisability of the findings to the larger category of automated journalism, which can include variants that use machine learning. Therefore, future investigations should analyse how variants of automated journalism that are more sophisticated technologically, such as those that use machine-learning-based generative AI, are integrated into the news production workflow, post-edited, and evaluated by readers. Furthermore, although the fitness of the multiple regression analysis model for both dependent variables (liking and perceived comprehensibility) was within an acceptable range, *r* square values show that a large part of the variance depends on other factors, the investigation of which lay outside the scope of this paper. Accordingly, future studies should explore which other narrative factors drive the liking and perceived comprehensibility of post-edited and automated news texts. Such findings would help to improve the automated production of data-driven news reporting for readers.

Declaration of conflicting interests

The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

ORCID iD

Sina Thäsler-Kordonouri https://orcid.org/0000-0003-3455-9315

Supplemental material

Supplemental material for this article is available online.

Note

1. The control variable's operationalisation in the survey can be found in the Supplemental Material.

References

Anderson CW (2012) Towards a sociology of computational and algorithmic journalism. *New Media & Society* 15(7): 1005–1021.

Beckett C and Yaseen M (2023) *Generating Change: A Global Survey of What News Organisations are Doing with AI*. London: Report for LSE's Journalism AI Project.

Borges-Rey E (2020) Towards an epistemology of data journalism in the devolved nations of the United Kingdom. *Journalism* 21(7): 915–932.

Boumans J, Trilling D, Vliegenhart R, et al. (2018) The agency makes the (online) news world go round. *International Journal of Communication* 12: 1768–1789.

Carlson M (2015) The robotic reporter: Automated journalism and the redefinition of labor, compositional forms, and journalistic authority. *Digital Journalism* 3(3): 416–431.

Caswell D (2019) Structured journalism and the semantic units of news. *Digital Journalism* 7(8): 1134–1156.

Caswell D and Dörr K (2019) Automating Complex news stories by capturing news events as data. *Journalism Practice* 13(8): 951–955.

Costera Meijer I (2020) What does the audience experience as valuable local journalism? Approaching local news quality from a user's perspective. In: Gulyas A and Baines D (eds) *The Routledge Companion to Local Media and Journalism*, Routledge, 357–367.

Danzon-Chambaud S (2023) Automated news in practice: A cross-national exploratory study. *Open Res Europe* 3: 95.

Danzon-Chambaud S and Cornia A (2021) Changing or reinforcing the "rules of the game". *Journalism Practice* 17(2): 174–188.

Diakopoulos N (2019) *Automating the News: How Algorithms Are Rewriting the Media*. Cambridge: Harvard University Press.

Diakopoulos N (2020) Computational news discovery: Towards design considerations for editorial orientation algorithms in journalism. *Digital Journalism* 8(7): 945–967.

Diakopoulos N and Koliska M (2017) Algorithmic transparency in the news Media. *Digital Journalism* 5(7): 809–828.

Dörr K (2019) Ethical approaches to computational journalism. In: Eldridge S and Franklin B (eds) *The Routledge Handbook of Developments in Digital Journalism Studies*. Oxfordshire, UK: Routledge, 313–323.

Dörr K and Hollnbuchner K (2017) Ethical challenges of algorithmic journalism. *Digital Journalism* 5(4): 404–419.

Franklin B (2007) *Local Journalism and Local Media. Making the Local News*. New York: Routledge.

Graefe A and 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.

Guzman AL and Lewis SC (2020) Artificial intelligence and communication. *New Media & Society* 22(1): 70–86.

Gynnild A (2014) Journalism innovation leads to innovation journalism: The impact of computational exploration on changing mindsets. *Journalism* 15(6): 713–730.

Haim M and Graefe A (2017) Automated news: Better than expected? *Digital Journalism* 5(8): 1044–1059.

Helberger N, van Drunen M, Moeller J, et al. (2022) Towards a normative perspective on journalistic AI. *Digital Journalism* 10(10): 1605–1626.

Hermida A and Young ML (2017) Finding the data unicorn: A hierarchy of hybridity in data and computational journalism. *Digital Journalism* 5(2): 159–176.

Jamil S (2023) Automated journalism and the freedom of Media. *Journalism Practice* 17(6): 1115–1138.

Koetsenruijter WM (2011) Using numbers in news increases story credibility. *Newspaper Research Journal* 32(2): 74–82.

Kuai J, Conill RF and Karlsson M (2022) AI ≥ journalism: How the Chinese copyright law protects tech Giants' AI innovations and disrupts the journalistic institution. *Digital Journalism* 10(10): 1893–1912.

Latour B (2005) *Reassembling the Social: An Introduction to Actor-Network-Theory*. Oxford, UK: Oxford University Press.

Leppänen L, Munezero M, Granroth-Wilding M, et al. (2017) Data-driven news generation for automated journalism. In: Proceedings of The 10th International Natural Language Generation conference, pp.188–197. Santiago de Compostela, Spain, 4–7, September 2017.

Lewis SC and Westlund O (2015) Actors, actants, audiences, and activities in cross-media news work. *Digital Journalism* 3(1): 19–37.

Milosavljević M and Vobič I (2019) Human still in the loop: Editors reconsider the ideals of professional journalism through automation. *Digital Journalism* 7(8): 1098–1116.

Møller LA, van Dalen A and Skovsgaard M (2025) A little of that human touch: How regular journalists redefine their expertise in the face of artificial intelligence. *Journalism Studies* 26(1): 84–100.

Porlezza C (2023) Promoting responsible AI: A European perspective on the governance of artificial intelligence in Media and journalism. *Communications* 48(3): 370–394.

Porlezza C and Di Salvo P (2020) Introduction: Hybrid journalism? Making sense of the field's dissolving boundaries. *Studies in Communication Sciences* 20(2): 205–209.

Preston CC and Colman AM (2000) Optimal number of response categories in rating scales. *Acta Psychologica* 104(1): 1–15.

Robinson JG (2019) *THE AUDIENCE IN THE MIND'S EYE: How Journalists Imagine Their Readers*. Tow Center for Digital Journalism A Tow/Knight Report.

Ruble RA (2017) Experimental manipulation. In: Allen M (eds) *The SAGE Encyclopedia of Communication Research Methods*. Thousand Oaks, USA: Sage Publications Ltd, 475–477.

Simon F (2022) Uneasy bedfellows: AI in the news, platform companies and the issue of journalistic autonomy. *Digital Journalism* 10(10): 1832–1854.

Simon F (2023) Escape me if you can: How AI reshapes news organisations' dependency on platform companies. *Digital Journalism* 12(2): 149–170.

Stalph F, Thurman N and Thäsler-Kordonouri S (2023) Exploring audience perceptions of, and preferences for, data-driven 'quantitative' journalism. *Journalism* 25(7): 1460–1480.

Sundar SS (1999) Exploring Receivers' criteria for perception of print and online news. *Journalism & Mass Communication Quarterly* 76(2): 373–386.

Tandoc EC Jr and Oh S-K (2017) Small departures, big continuities? *Journalism Studies* 18(8): 997–1015.

Tandoc EC Jr, Wu S, Tan J, et al. (2022) What is (automated) news? A content analysis of algorithm-written news articles. *Media & Jornalismo* 22(41): 103–120.

Thäsler-Kordonouri S (2024) What comes after the algorithm? An investigation of journalists' post-editing of automated news text. *Journalism Practice* 0(0): 1–20. https://doi.org/10.1080/17512786.2024.2404692.

Thäsler-Kordonouri S and Barling K (2023) Automated journalism in UK local newsrooms: Attitudes, integration, impact. *Journalism Practice* 19(1): 58–75.

Thäsler-Kordonouri S, Thurman N, Schwertberger U, et al. (2024) Not descriptive enough and too many numbers: Why readers find data-driven news articles produced with automation harder to understand. *Journalism* 0(0): 1–21. DOI: 10.1177/14648849241262204.

Thurman N (2019) Computational Journalism. In: Wahl-Jorgensen K and Hanitzsch T (eds) *The Handbook of Journalism Studies, Second Edition*. New York: Routledge.

Thurman N, Dörr K and Kunert J (2017) When reporters get hands-on with robo-writing. *Digital Journalism* 5(10): 1240–1259.

Thurman N, Stares S and Koliska M (2024) Audience evaluations of news videos made with various levels of automation. *Journalism* 26(1): 3–23.

Waddell TF (2019) Attribution practices for the man–machine marriage: How perceived human intervention, automation metaphors, and byline location affect the perceived bias and credibility of purportedly automated content. *Journalism Practice* 13(10): 1255–1272.

Wang S and Huang G (2024) The impact of machine authorship on news audience perceptions: A meta-analysis of experimental studies. *Communication Research* 51(7): 815–842.

Willis GB (2016) Questionnaire pretesting. In: Wolf C, Joye D, Smith TW and Fu Y (eds) *The SAGE Handbook of Survey Methodology*. Thousand Oaks, USA: Sage Publications Ltd, 359–381.

Wilner T, Muse H, Coddington M, et al. (2024) The Research-Practice Gap in Journalism: Why It Exists and How We Can Address It. (White paper). *Journalism Bridging Project and American Press Institute* https://www.journalismbridgingproject.com/.