



UvA-DARE (Digital Academic Repository)

News in perspective

What is news to news users in the digital age?

Lin, Z.

Publication date

2025

[Link to publication](#)

Citation for published version (APA):

Lin, Z. (2025). *News in perspective: What is news to news users in the digital age?* [Thesis, fully internal, Universiteit van Amsterdam].

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

Chapter 5

Conclusion and discussion

Conclusion and discussion

In today's era, marked by emerging technologies such as social media, smart devices, and artificial intelligence, the landscape of news has undergone significant changes. This new environment has three dynamically intertwined components: news *content* with blurry genre demarcation, the thriving digital *platforms* that serve both news creators and consumers, and the widespread application of *algorithms* in content curation. In this contemporary setting, individuals can access news information online from anywhere, at any moment, and this naturally reshapes their perception of what constitutes news.

This dissertation therefore asked the central question: “*What is news to news users in the digital age?*”. This question was unpacked and divided into three research questions, each corresponding to an empirical chapter. Innovative data collection approaches and computational methods were used to address these research questions, together shedding light on the evolving digital news landscape. Throughout these studies, the emphasis was on news perception from the users' perspectives. This final chapter discusses the main findings and conclusions, answering these three research questions, and together the central question. Additionally, it addresses the theoretical, methodological, and practical implications of this dissertation and concludes with some final remarks.

Main findings

RQ1: How can news, as an evolving genre with blurry boundaries, be measured along abstract dimensions, such as genre cues?

Chapter 2 addressed this question by organizing news-related items, from news articles and opinion pieces to YouTube videos, podcasts, satirical programs, and talk shows, on a two-dimensional grid consisting of the dimensions *Factuality* and *Formality*. The position of each item was based on the density of formal and factual sentences within its content. These dimensions were conceptualized as genre cues, the signals that users pick up to recognize an item as belonging to a specific genre,

in this case, news. Supervised machine learning was used for such large-scale classification tasks. Trained with two annotated corpora ($N = 18,305$) created by ourselves, fine-tuned large language models (i.e., BERTje) outperformed baseline models, including Naïve Bayes, logistic regression, and support vector classification, when classifying sentences in both the *Factuality* and *Formality* task. The models can be applied to news-related items in a broader sense, as long as the textual representations available. This was also demonstrated in the two showcases in **Chapter 2** as well as in the analyses in **Chapters 3** and **4**. Generally speaking, while news-related content is diverging from the traditional format, which is almost entirely factual and formal in linguistic style, it does not stray far from these standards. Most items studied in this dissertation still appeared to be recognized as rather formal and factual, except for the podcast content, which appeared to be more opinionated and informal. Such patterns are also in line with one's expectation, providing further evidence that our machine learning approach is valid and reliable.

RQ2: To what extent do different definitions of news lead to differences in users' news repertoires on digital platforms?

To answer this question, two batches of YouTube watching histories were collected in **Chapter 3** through data donations. The first one was online via a panel company reaching a (roughly) representative Dutch sample ($N_{participants} = 344$, $N_{items} = 3,267,244$), while the second batch was collected offline at a music festival targeting a younger population ($N_{participants} = 288$, $N_{items} = 3,103,615$). Before donating the data, participants were asked to indicate news-related channels from their own YouTube histories, which later allowed us to compare the traditionally-/professionally-defined news repertoires and the user-defined news repertoires (either on the individual or aggregated levels). YouTube video subtitles were analyzed using large language models (i.e., BERTje and RoBERTa) in terms of classic journalistic dimensions, namely *Factuality* and *Formality*. Results showed that users have varying definitions of what constitutes news, covering a broader range of content than the traditional/ professional definition. Furthermore, using different definitions significantly shaped our understanding of news repertoires on YouTube. Adhering to the traditional/ professional definition resulted in a more factual and formal news repertoire of the same users, compared to using the users' definitions. This effect hold true across different languages (i.e., Dutch and English)

and samples. Moreover, it is moderated by individual characteristics, for example, demographics (e.g., education level) and political traits (e.g., political leaning, political interest, and political efficacy). For instance, those with lower educational levels or political efficacy hold a definition of news that deviates further away from the traditional one, leading to larger differences in the resulting news repertoires with different news definitions applied.

RQ3: How is perceived news-ness influenced by textual and contextual genre cues?

In order to address this question, an online field experiment was conducted in **Chapter 4** for over two weeks on a customized news application with a live news feed ($N_{items} = 11,040$ from eight diverse sources in the Netherlands, retrieved during the experiment). It adopted a between-subject design with two conditions differing on whether the news feed was generated by algorithmic recommendation ($N_{random} = 60$ versus $N_{recommendation} = 54$). During the experiment, participants were shown nine news-related items each time they refreshed the page, with the central one as the “surprise box” concealing a randomly selected item to facilitate serendipitous news discoveries for both groups. Participants were asked to rate the news-ness of the items they clicked on ($N_{participants} = 114$, $N_{items} = 4,342$). It is worth noting that participants in the recommendation conditions were aware of the monotonous nature of their news feed and exhibited a significantly higher level of serendipity seeking, compared to those in the random condition. In terms of news-ness, observational data indicated that the degree to which an item was perceived as high in news-ness by users is largely dependent on text-based genre cues, specifically its *Factuality* and *Formality* predicted by the models trained in **Chapter 2**. Consistent with **Chapter 3**, this relationship is influenced by individual characteristics, such as age, education level, and political leaning. For instance, those with a lower education level tend to be more lenient and perceive informal content with higher new-ness, compared to the higher educated ones. Additionally, user perception is influenced by contextual genre cues, namely the item’s recommendation nature (whether it is curated by recommendation or at random) and presentation style.

Theoretical implications and future research

This dissertation has important theoretical implications on news from the users' perspective. It reflects on how we understand news perception and news consumption in the digital age, moving beyond the previous mostly supply-centric literature.

A multidimensional framework for studying news

This dissertation introduced a multidimensional framework that contributes to the conceptualization of news from various perspectives. This framework was proposed and showcased in **Chapter 2** for analysing news content, while put into application in **Chapter 4** for mapping news users based on the news content they clicked on during the online field experiment. As a new approach to news genres, it emphasizes the evolving forms and narrative styles of news, complementing the traditional focus on news values that has dominated communication science research (Broersma, 2007; Dent, 2008). By incorporating dimensions like *Factuality* and *Formality*, it moves beyond the restrictive genre categorizations (Buozis & Creech, 2018), more accurately capturing the dynamic and fluid nature of news genres (Steensen, 2011). Furthermore, by conceptualizing *Factuality* and *Formality* as genre cues that users can pick up and use for genre recognition, this approach allows valuable insights into news perceptions to be provided (Costera Meijer, 2020).

While the current emphasis on *Factuality* and *Formality* is grounded in their strong ties to classic journalistic principles, future research is encouraged to explore additional dimensions. These could include the density of humor to identify satirical content (Otto et al., 2017), or language complexity measures such as readability (Graefe et al., 2018; Pérez-Rosas et al., 2017).

What is news to news users

Following the call of taking a (*radical*) *audience turn* (Costera Meijer, 2007; Swart et al., 2022), this dissertation encourages researchers to rethink the “*what is news*” question. Moving beyond the classic discussion from the supply-side of journalism evaluating news values and professional standards (Gans, 1979; Harcup & O’Neill, 2017), this dissertation provided evidence from the users’ perspective in news perception, along with large-scale user data on news consumption.

It is increasingly acknowledged that the definition of news is becoming blurry (Edgerly & Vraga, 2020), particularly on digital platforms where diverse sources and genres blend in visually engaging formats (Lewis, 2020). **Chapter 3** substantiated this argument with quantitative evidence, revealing that differing definitions of news significantly influences our understanding of users' news repertoires. The findings suggested that repertoires shaped by user-defined definitions tend to be more informal and opinionated compared to those constructed based on professional standards. This resonates with the suggestion of taking an *audience turn* (Costera Meijer, 2020; Swart et al., 2022), emphasizing the importance of adopting a user-centric perspective in news studies. Future research is encouraged to reconsider the traditional approaches to studying news consumption. While analyzing news using professional definitions holds value, it is important to understand how users interact with and adapt to the evolving news landscape. Furthermore, in exploring topics such as the effects of news, it becomes essential to examine news consumption based on the users' own definitions. Additionally, the relationship between news definitions and resulting repertoires varies across different user groups. Future research is therefore advised to focus on specific populations when studying news, such as young people (Swart & Broersma, 2023), and consider other latent concepts that may influence how individuals perceive and interact with news like politically-charged variables (Eddy, 2022).

Chapter 4 extended the investigation of news perception and specifically look into the genre cues users recognized to decide the extent to which certain content is considered news. It identified a variety of genre cues related to both content and context upon which users rely to evaluate the news-ness of content. Users mainly turn to formal and factual linguistic features as indicators for news content, underscoring the lasting relevance of traditional journalistic principles like objectivity and professionalism (Harcup & O'Neill, 2017). Consistent with prior findings on news definitions and user characteristics in **Chapter 3**, this chapter also noted variability in reliance on these textual cues among users, such as those with differing education levels and political orientations. Future studies are encouraged to expand the examination to other possible textual genre cues and user characteristics that may influence news perception.

Moreover, the recommended nature of an item influences users' perceptions of news-ness. It is possible the recommendation could act as both a content cue, where

familiarity leads to positive feedback for consistent content (Wieland et al., 2021), and a contextual cue, where users' attitudes towards algorithmic recommendations affect their evaluations (Thurman et al., 2019). Results showed that the status of an item as recommended or not significantly interacts with textual cues, enhancing or reducing the perceived news-ness based on a certain threshold. Content deemed roughly 75% news-like benefited from recommendation endorsement, while content below this threshold was perceived as even less news-like compared to content delivered randomly. These insights suggested that recommendations play a role in shaping news perception and call for a more detailed examination of these dynamics. In addition, **Chapter 4** also revealed the crucial role of presentation as a contextual cue in shaping how news is perceived. Items featured through the "surprise box", an interface designed to introduce elements of serendipity, were generally viewed as less news-like, irrespective of their textual cues or user demographics (Abdollahpouri et al., 2021). These findings demonstrate that presentation is just as important as content (Abdollahpouri et al., 2021), indicating that unconventional online news formats might struggle for recognition if users are accustomed to traditional styles (Broersma & Harbers, 2018). Conversely, the risk arises that if false information adopts these familiar news formats, users might more easily mistake such content for high-quality news.

Collectively, these findings explored the dynamics of different genre cues in shaping individuals' perceptions of news. Despite individual variations, users tend to lean on traditional criteria for recognizing news, with a focus on *how* it is communicated, including writing style, delivery mechanism, and presentation format. In an information-saturated age, employing these genre cues for distinguishing news content offers both benefits and challenges. From the user's perspective, such cues allow for easy identification of legitimate and pertinent news content, targeting those resembling traditional news presentation. However, this strategy may also leave users susceptible to mis-/disinformation mimicking conventional news formats. For news providers, using genre cues in news communication reflects the ongoing adaptation to digital platform technological affordances (Nieborg & Poell, 2018). Users clinging to traditional standards may counter journalists' efforts to innovate news communication on digital platforms. Future research is encouraged to explore other genre cues and examine exactly whether and how they contributed to users' news perception.

The active role of news users in the recommendation context

The findings in **Chapter 4** demonstrated that by design, the diversity of news consumption via highly personalized recommender systems is typically lower than that observed with random news selection (e.g., Nguyen, Hui, Harper, Terveen, & Konstan, 2014). However, this does not inevitably lead to negative feedback loops where users would be trapped with a less and less diverse news diet. Observations of news users' perceived diversity revealed that those in the recommendation group were aware of their limited exposure. More importantly, the introduction of a "surprise box" feature intrigued users across different demographic and political backgrounds.

Compared to the random condition, individuals in the recommendation group more actively sought a break from their news feeds, which they may perceive as too monotonous, by engaging with the "surprise box". This finding highlighted users' discontent with (too) similar content and their active role in escaping the narrow feedback loop, paving the way towards diverse news consumption. On the one hand, it was confirmed that algorithmic accuracy alone does not guarantee a positive user experience and may, in fact, diminish it over time (Dubois & Blank, 2018; Fletcher & Nielsen, 2018; Möller et al., 2018). On the other hand, this challenged previous assumptions that recommendation systems inherently lead to polarization (e.g., Pariser, 2011; Stroud, 2010), demonstrating the potential of interface designs to encourage serendipity (Smets et al., 2022), and – again – showing that users are not merely passive recipients of algorithmically curated news. Future research should explore other interface features that promote serendipitous encounters and thoroughly assess their effects on news perception and news consumption, particularly within the context of news recommendation systems.

Methodological implications

The methodological challenges in studying news consumption in the digital age lie in both collecting and analyzing a large amount of user data in a valid and reliable manner. This dissertation presented innovative approaches aiming towards this goal. The next paragraphs will discuss the methodological implications.

Collecting digital trace data with users

Traditional methods for collecting media consumption data have predominantly relied on self-reports, such as surveys, diaries, and interviews. Despite fruitful findings gathered, the accuracy of self-report data is often compromised by the difficulty participants may have when recalling their past usage or their inclination to over report and provide socially desirable answers (Prior, 2009b). To address these limitations, this dissertation collected digital trace data for measuring news consumption: In **Chapter 3** participants donated their YouTube watching histories directly from Google Takeout, while **Chapter 4** documented every click participants make on the news application during the experiment period lasting for approximately two weeks. This approach not only offers more accurate and detailed records of news consumption behaviors but also enables the combination of user data with the users' own survey answers. In both chapters, these survey questions proved useful for investigating how users interpret their own data on a large scale, going beyond mere demographic and common attitudinal information. Specifically, in **Chapter 3**, users' donated data were used to ask personalized survey questions about their news consumption, while in **Chapter 4**, users answered questions about the articles that they clicked on.

Furthermore, this dissertation emphasized the importance of collecting digital traces with users, listening to their input during the data donation process. One key strategy is the inclusion of qualitative questions. For example, in **Chapter 3**, users who failed to complete the data donation process were asked about reasons for their withdrawn, which highlighted concerns over privacy, the effort required, and technical difficulties encountered. This feedback underscored transparency and clear instructions to reduce the high dropout rate in data donations, which were later achieved during offline data collection at a music festival. In this setting, lab personnel accompanied participants, guiding them step-by-step while having a conversation with them that offered insights beyond the user data itself. These discussions revealed participants' comfort with data donation, their reactions to their past media usage, and their thoughts on algorithmic recommendations. In this way, collecting data together with the users addressed most challenges concerning the high drop-out rate. Similarly, in **Chapter 4**, feedback from users about their experience with the news application provided valuable insights into not only app improvement, but also their preferences for news diversity and the opt-out feature

(i.e., the “surprise box”) when reaching saturation from recommendations.

Another suggestion for future researchers is to not overlook the benefits of non-financial incentives. While monetary rewards are common motivators, this dissertation highlighted the effectiveness of alternative rewards. In **Chapter 3**, visualizations of participants’ digital footprints not only educated them about what kind of data tech companies are collecting about them, but also made the data donation process more transparent and engaging. At the music festival, being able to participate in groups and learn more about themselves together with friends and families contributed to a fun experience. Together with other fun elements like printing tattoos of participants’ favorite music genres, based on their YouTube data, the potential of non-financial incentives to encourage participation in research requiring more time, effort, and sensitive data from the participants is showcased. When there is no monetary rewards involved, it can still be a successful data collection if researchers are collecting data with the users and making it a fun experience. Feedback from **Chapter 4** suggested that participants appreciate accessing live news information without paywalls or ads, which becomes their reasons for continuing using the application even after the experiment.

To be able to meaningfully interpret user data, understanding their perceptions of their own digital traces is key. This dissertation, particularly in **Chapters 3** and **4**, demonstrated how to incorporate users’ perspectives into the analysis of their own consumption data. In **Chapter 3**, participants were presented with their most-visited YouTube channels and asked to indicate which they consider as news related, with the option of adding others not listed. This threshold at the channel level helped identify all consumed videos that are deemed news related by the users themselves. **Chapter 4** took a similar approach, asking participants during the experiment to evaluate the news items they click on, rating them on their relevance and on the extent to which they are viewed as news. By doing so, the concept of news-ness was measured at the individual item level. Involving users in identifying what they consider news not only aligns with the conceptual argument of this dissertation but also provides a practical method to efficiently filter relevant data from an overwhelming amount of digital traces. When studying news consumption, the analysis should focus on content that users find meaningful, rather than merely tracking their clicks (Verbeij et al., 2021).

This strategy also serves to evaluate the validity of digital trace data, which may

not fully capture users' past behaviors due to shared accounts, disabled tracking features, or technical errors. Before donating their data, in **Chapter 3** users were asked to assess its accuracy and whether it belongs to themselves only, offering face validation of the data. This dissertation advises future research on digital traces to integrate qualitative insights in data analysis, whether to evaluate data accuracy, obtain valid and reliable measurements, or both, enhancing the understanding of digital consumption through a more nuanced, user-centric perspective.

Data analysis with large language models

All chapters in this dissertation demonstrated the opportunities of using computational methods, particularly large language models (LLMs), for news analysis. These models have shown better performance than traditional machine learning models in tasks like sentence classification and punctuation prediction due to their superior contextual understanding of the text, which allows them to handle linguistic ambiguity and complexity in a better way (Devlin et al., 2018). This capability is particularly beneficial for researching theoretically fuzzy concepts such as news, for which explicit rules are often challenging to define yet there are indeed patterns that can be learned from observations (Boumans & Trilling, 2018). Given the fast and fruitful development of large language models, future studies in communication science, are encouraged to use that for their benefits, with the following recommendations in mind.

First, rather than directly using the pre-trained models, it is advisable to tailor the models for specific research objectives by fine-tuning them with domain-specific training data and exploring various parameter settings in order to discover the optimal configuration. For instance, this dissertation tailored a model specifically to predict formality within the news genre, differentiating from common practices in the computational field where news content is directly labeled as "formal" when used as training data for formality tasks (Sheikha & Inkpen, 2010). An important step is to create a diverse and representative corpus for the tasks, enabling the model to more effectively transfer its pre-trained knowledge. Specifically for news-related tasks, it is now feasible to include diverse news formats, such as spoken news content, by using automatic transcription through speech-to-text models, followed by manual corrections. This approach addresses the previous challenge of studying large-scale spoken news, which was often limited by the lack of available transcripts (Chelsey, 2021). Moreover, in building such corpora, there

are also useful tips for annotation, such as employing majority voting or labeling on a scale that can later be grouped into a binary format, allowing for more nuances and eventually a more refined dataset.

The second piece of advice emphasizes the importance of validation when using large language models, especially when applying them directly. Incorporating a small set of annotated data can prove beneficial for this purpose. Additionally, performing tasks across multiple languages can further validate the results, enhancing their robustness.

Practical implications

This dissertation also offers practical implications to not only news makers but also society in general.

News perceptions benefit news experiences

In today's new media landscape, newsrooms have been striving to adapt to digital technologies and emerging platforms. Their journalistic decisions benefit from constant user feedback. Tracking user metrics and making efforts to engage the users have become common practices (Kristensen, 2023). With the acknowledgment that newsrooms may no longer be at the center of news-related activities, this dissertation established the relevance of understanding news perceptions from news users. These user insights, in addition to traditional click data, enable journalistic practitioners to more deeply understand, for example, not just superficially what content attracts traffic, but also why certain content resonates with their users. To obtain such perceptions, news makers are encouraged to maintain constant and direct dialogues with their users, either in the format of interviews or simply in an active comment section on certain digital platforms (e.g., on YouTube, *NOS op 3* frequently reacted on users' comments regarding the current video or recommendations on future content). When passive consumers turning into active participants, a better understanding of them would allow news makers to effectively improve their future news offerings, thus not only attracting their current users but also new users. Ultimately, user perceptions can contribute to a more pleasant news experience for both news makers and users, and in the long run, a more informed and participatory society.

Education on digital literacy should be emphasized

The findings of this dissertation highlighted the importance of digital literacy education, especially (fake) news literacy (Tamboer, Vlaanderen, Bevelander, & Kleemans, 2024). As demonstrated in **Chapter 4**, news users to a great extent rely on genre cues such as *Factuality* and *Formality* to differentiate between news and non-news. However, relying on such shortcuts during information gathering can be problematic. These attributes are easily emulated in mis-/disinformation, or so-called “fake news”, leaving users vulnerable to misleading and deception (Bennett & Livingston, 2018; Lazer et al., 2018; Waisbord, 2015). In today’s environment where this kind of false information prevails and often resembles the format of traditional news, it is vital to equip users with the skills to discern news of quality.

Moreover, this dissertation revealed the need to raise users’ awareness about their personal data rights. Most users, as documented in **Chapter 3**, are insufficiently familiar with GDPR or other data protection laws. When learning how to access their own data, many expressed surprise at the extent of their information held by big tech companies like Google. Such awareness is crucial because citizens can only exercise their rights and protect their privacy, for instance, by disabling certain tracking features, when they are familiar with their data rights and know how to access their personal data. Therefore, education on digital literacy should not only help news users in recognizing trustworthy information but also with their data rights, both of which empower news users and contribute to building a healthy media environment.

Towards designing better recommender systems

This dissertation also shed light on news users’ expectations on how recommendation systems should be designed on digital platforms. Feedback from users in **Chapter 4** highlighted their preferences for greater transparency, autonomy, and diversity in news recommendations, suggesting that achieving a satisfactory user experience extends beyond merely meeting algorithmic accuracy (Smets et al., 2022). After all, the path to achieving a positive user experience aligns closely with designing recommendation systems that uphold the democratic benefits of news, bridging the goals of a recommender system designer and a communication researcher. This dissertation emphasized the merits of respecting the active role of news users, who show appreciation of a positive feedback loop

towards diversity. This insight points to a promising direction for future development in news recommendations, one that prioritizes user expectations and democratic benefits in its design.

Concluding remarks

In conclusion, this dissertation makes contributions to the field of (digital) journalism, both theoretically and methodologically. Theoretically, it introduced a multi-dimensional framework for conceptualizing news on a broad and expansive scale, moving beyond traditional discrete genre classifications. It advocated rethinking the fundamental question, “what is news”, building on the traditional discussions about news values and professional standards from the newsrooms’ perspective. The empirical chapters deepened the understanding from the users’ perspective by integrating their own news perceptions with large-scale user data. Moreover, the findings shed new light on the users’ preference for diversity in news recommendations and their active role in seeking serendipitous news encounters over monotonous feeds. Methodologically, this dissertation used innovative data collection approaches for collecting digital trace data, including data donations applicable in both online and offline contexts and online field experiments. It also demonstrated the advantages of computational methods in news analysis, applying large language models that showcase to be outperforming traditional machine learning models.

Taken together, this dissertation highlighted the relevance of a user-centric focus in (digital) journalism research, providing newsrooms with valuable user insights into engaging journalistic experiences with democratic benefits. Furthermore, it emphasized the importance of digital literacy education, enabling users to critically discern news. All in all, these efforts contribute to fostering a well-informed digital society, echoing the evolving news perception and news consumption in the digital age.