Telling a Different Story: A Longitudinal Investigation of News Diversity in Four Countries

Erik de Vries, Rens Vliegenthart & Stefaan Walgrave

View supplementary material

Published online: 01 Sep 2022.

Submit your article to this journal

Article views: 415

View related articles

View Crossmark data

Telling a Different Story: A Longitudinal Investigation of News Diversity in Four Countries

Erik de Vries [a], Rens Vliegenthart [b] and Stefaan Walgrave [c]

aDepartment of Media and Social Sciences, University of Stavanger, Stavanger, Norway; bStrategic Communication Group, Wageningen University & Research, Wageningen, the Netherlands; cDepartment of Political Science, University of Antwerp, Antwerp, Belgium

ABSTRACT

News diversity is an important concern of journalism scholars, as its presence or absence can have a profound effect on democratic debate and the information available to citizens. Many have speculated that news diversity decreases over time, due to changing economic circumstances. This expectation especially applies to newspapers. Using nearly two decades of newspaper data from four European countries (Denmark, The Netherlands, Norway, UK), we do not find this expected decrease in news diversity. When conducting pairwise, automated comparisons between articles published on the same day in the same country, we rather find a modest over time increase in diversity between newspapers. This result suggests that newspapers differentiate rather than converge in the content they offer, shedding a more positive light on the evolution of the press in our current high-choice media environments.

Introduction

During the last decades, worries have grown about a potentially increasing convergence of news coverage in traditional media. If news media indeed make increasingly similar news selection and framing choices, this could be considered worrying from a democratic perspective. External news diversity, with news outlets competing by offering different stories, is considered by many scholars to be an important feature of a healthy information environment. When different outlets highlight different stories, different elements of the same story, or evaluate the story differently, this generates a rich and more pluralist context to spark democratic debate. It also allows citizens, if they want, to be confronted with different opinions and form themselves a nuanced idea about the facts and how to evaluate them.

The reasons for widespread pessimism regarding decreasing external news diversity are strongly related to structural changes in the news business. Two complementary mechanisms are supposed to be at work. First, increasing media competition, audience fragmentation and hybridization of news lead to more pressure on journalists to produce more news stories. Without proper time to select deviant news stories and to develop their own approach, journalists produce news stories that are interchangeable

and that are heavily affected by the information subsidies provided by the story's stakeholders. Second, economic pressure on the news sector leads to a concentration of different news outlets in the hands of fewer owners who push for increased collaboration and integration of the newsrooms in their portfolio. This is likely to almost mechanically generate an increasing overlap in news stories. The two causes of the alleged decrease in diversity are well-documented; journalist surveys almost invariably point to increased productivity (Hanusch 2015 in Australia; Jyrkiäinen and Heinonen 2012 in Finland; Raeymaeckers, Paulussen, and De Keyser 2012 in Belgium) and news outlets are increasingly concentrated in conglomerates. Even so, actual empirical proof of their alleged effect on news diversity is rare, and the few studies that exist present mixed evidence.

In this paper, we contribute to the ongoing debate about news diversity. Our study is empirical, not normative. We do not take a normative position and do not claim that decreasing news diversity is invariably a bad thing; in fact, some studies show that news, if diverse, is less used by the consumers (Van Aelst et al. 2017) and concentrated attention could under certain circumstances spark societal debate and put pressure on decision makers to be responsive (Walgrave et al. 2017). What we do here is empirically examining news diversity in a range of countries for a long time period. The content of individual newspaper articles published on the same day, in four countries (Denmark, The Netherlands, Norway and the UK) is analyzed and compared. This is done for three newspapers per country, over a period of twenty years (2000–2019). While newspapers all over the globe have suffered serious declines in readership figures, they are still among the most frequently used media sources in the countries under study (Newman et al. 2021) and have for example shown to exert considerable political agenda setting power (Langer and Gruber 2021). Additionally, in times of dynamic and fastly changing media landscapes, they have been among the few outlets that have sustained a prominent position and consequently offer ample opportunities for systematic over-time comparisons. In total, 6 million newspaper articles are examined. We rely on an automated procedure to identify for each day and each newspaper pair in a country which articles are most alike. These most similar article pairs (between newspapers) are also most likely to deal with the same topic. If the percentage of article pairs about the same topic goes up over time, we argue that there is a tendency towards less topic diversity. In addition, we leverage a computer-generated sentiment dictionary to examine the degree to which article pairs about the same topic also present these topics in a similar sentiment context. This allows us to determine if there is a trend towards less sentiment diversity as well. Our results do not support the expectation of less diversity, however. In none of the four countries, we find decreasing topic diversity. The same applies to sentiment diversity: it does not decrease over time. If anything, in some countries, we see very modest signs of the exact opposite pattern, namely that newspaper content becomes more diverse over time. We discuss this contra-intuitive trend and link it to the further growth of interpretative journalism, whereby newspapers try to differentiate themselves in an increasingly competitive environment by offering distinct news facts and interpretations.

Why News Diversity Would be Decreasing

The importance of news diversity is a topic of societal and scientific debate. Many assume that, in one way or another, diversity is relevant for the functioning of democracy (Beckers

et al. 2019; Napoli 1999; Sjøvaag 2016; Vogler, Udris, and Eisenegger 2020). Western democracies rely on media and their supply of information to function properly. Voters need to be able to obtain diverse information in order to be able to cast an informed vote. Despite the fast-changing information environment, legacy mass media remain the most prominent source of information for the majority of citizens in most countries. At the same time, politicians too need access to media and the information media provide to communicate with voters about relevant issues (Van Aelst and Walgrave 2016). There is debate about how large news diversity should be exactly, and some argue that too much diversity is not good either, as it leads to audience fragmentation and the public sphere falling apart (Roessler 2007). Van Cuilenburg (1999, 199) states that diversity in the media cannot be evaluated in the abstract, as it " … should always be compared with relevant variations in society and social reality." In other words, media diversity should ideally be a reflection of the actual diversity of ideas and opinions in society (see also Joris et al. 2020).

Since measuring the real diversity of opinions and ideas in a society is hardly possible, it is problematic to observe the relative diversity in media markets—that is the diversity relative to real-world differences. Most research, therefore, looks at absolute diversity, focusing in particular on over-time and cross-context variation. Some scholars have for example examined the absolute diversity of topics in news stories from different media. Do they cover the same events during a particular news cycle (see Boczkowski and Santos 2007; Joris et al. 2020)? Topic diversity is only part of the story, though. News media that are supposed to represent the diversity of ideas and opinions in a society should also to some extent employ different interpretations when telling their stories. In fact, while investigating news topics provides an idea of what news media report about, it does not tell anything about the "interpretation, evaluation and/or solution" (Entman 2003, 417) related to an event. In line with those interpretations, in this paper, we consider external news diversity as the extent to which, in a given period of time, different news outlets (1) report on different topics (2), and when they report on the same topics, to what extent they use a similar tone or sentiment. Note that this definition explicitly excludes internal news diversity or the diversity of content features within a news outlet. In the remainder of this text, the term "diversity" is exclusively used in the context of external news diversity. Additionally, we constrain ourselves to topics and sentiment and do not single out events or more elaborate frames.

Researchers have speculated that, in many countries, diversity of the news has diminished (Lee 2007; Schudson 2011). The dominant account holds that due to underlying structural economic evolutions, news corporations have been forced to change their strategy, with decreasing news diversity as a consequence. This so-called "newspaper crisis" is characterized by high levels of competition in shrinking markets (Curran 2010). Under these circumstances, newspapers all across Europe have been struggling to keep their businesses profitable (Brüggemann, Esser, and Humprecht 2012; Lewis, Williams, and Franklin 2008; Vogler, Udris, and Eisenegger 2020). For a large part, this is due to people switching from newspapers to other news sources, especially on the Internet or social media. Newspapers have suffered from declining readership numbers, and lower advertising and subscription revenues. There exist different and contradictory accounts of how such conditions of high competition affects newspapers' strategy. Hotelling's Law (1929) posits that under conditions of high competition competitors generally

tend to compete on price, rather than on product differentiation and quality (see also Van Cuilenburg 1999). High competition in the news market would thus lead to cost cutbacks and, hence, to less news diversity. At the same time, some studies found the exact opposite, being that newspapers invest more in product quality, and thus product diversity, when the pressure of competition increases (Lacy and Simon 1993).

Still, the less diversity argument is more frequently present in the literature because there is more proof of the fact that high competition has led to cost reduction and staff cuts. This is the first likely cause for a decrease in diversity: less and less journalists must produce the same amount of content. Taking some time to develop a story, and thereby make it different from that of a competing outlet, is therefore often not possible. Indeed, more resources for reporting and more specialized journalism lead to more diverse news both in terms of the events covered and the evaluation. Or, inversely, a higher workload leads to journalists increasingly relying on content—often called "information subsidies" (Gandy 1980)—produced by external sources, such as PR and news agencies (Boumans et al. 2018; Vogler, Udris, and Eisenegger 2020). Under time pressure, journalists from different media outlets cannot but rely on the same external sources, leading to a reduction in news diversity.

Yet, even if staff were not cut, the news business has changed rapidly and news outlets such as newspapers are not only present on the print news market but have become broad news providers with elaborate news websites and a strong presence on social media. This entails a 24/7 online news presence, meaning that journalists simply have to produce more news on each day (Paulussen 2012). The effects thereof are comparable to the effect of staff cuts and leads to a second reason for declining news diversity: less time per news item and, hence, less chance to develop a different topic and angle choice.

Besides staff cuts and increased workload, a third reason for the alleged decline in news diversity is that publishing houses have merged as yet another consequence of the newspaper crisis (Curran 2010). The result is a concentration of ownership, with only a few publishing houses owning the majority of national newspapers in most European countries (Picard 2014). This mechanism has, according to some, also led to an overall decrease in news diversity. If newsrooms lose autonomy and are forced to (partly) pool resources with other newsrooms, the consequence can only be that the news choices and output of the collaborating outlets become more similar (Beckers et al. 2019; Dailey, Demo, and Spillman 2005; Hendrickx and Ranaivoson 2019). So, trends in media concentration and news outlet ownership have arguably diminished news diversity (Baker 2007).

Topic and Sentiment Diversity

Notwithstanding the widely shared pessimism regarding decreasing news diversity, the number of empirical studies that demonstrates decreasing diversity over time remains very small and their findings are mixed. Joris et al. (2020) provide a systematic review of news diversity studies. These studies differ in the way the conceptualize and operationalize news diversity, as well as in scope of the investigation, and maybe not surprisingly, in the results. Regarding conceptualization, the literature overview by Joris and colleagues (2020) finds thirteen studies that have looked at topic diversity (almost none with over time comparisons) but hardly any study in their overview considers, for

instance, viewpoint diversity (e.g., Day and Golan 2005 who looked at viewpoint diversity in op-ed contributions within the same newspaper; see also Rodgers, Thorson, and Antecol 2000) or actor diversity (e.g., Masini and Van Aelst 2017). As mentioned earlier, the variation between outlets in the interpretation and evaluation of stories may, from a democratic diversity perspective, be even more important than whether they actually cover the same events. Indeed, work in political communication has shown that how journalists discuss a topic or event matters for how news consumers digest it. Sentiment, either attributed directly towards specific objects, or at the more general level of a news item, is a frequently considered content feature (see Boukes et al. 2020). In general, it has been shown that negative coverage has a stronger effect on the audience than positive coverage (Soroka and McAdams 2015; Vliegenthart et al. 2021) and consequently is more widely used to attract the largest audience possible (Damstra and De Swert 2020). For example, if the economy is generally discussed in negative terms, this yields lower levels of consumer confidence, and those effects are larger than for positive coverage. Similar to topic diversity, general sentiment in issue coverage can differ in its level of diversity. If then competition for readers increases, it might well be that variety in sentiment by which topics are discussed decreases as well, following a similar logic as for topic diversity.

Only a few longitudinal studies exist. They indeed show that newspaper coverage becomes less diverse (Boczkowski and Santos 2007 in Argentina; Vogler, Udris, and Eisenegger 2020 in Switzerland) and more reliant on the same external sources (Vogler, Udris, and Eisenegger 2020 in Switzerland). Other studies do not find a decrease in diversity (Beckers et al. 2019 in Belgium). Yet, these few longitudinal studies all look at one specific country, they examine different time periods, and their measurements of diversity vary largely. Hence, although the underlying economic trends of competition, increasing work pressure, staff cuts and mergers are well-established, there is no firm proof for the often-assumed consequential decreasing news diversity. Further, all extant longitudinal studies only look at topic diversity and none take other aspects of diversity into account. Our study tries to improve on previous work by (1) testing the generalizability of trends by looking at newspapers in four different countries (Denmark, Norway, UK and the Netherlands) belonging to different media system types (Hallin and Mancini 2004), (2) by covering a long time period (2000–2019), (3) by including all types of hard news content, (4) by drawing on a systematic and reliable automated approach to assess diversity (Amsalem et al. 2020; Vogler, Udris, and Eisenegger 2020), and (5) by looking both at both topic diversity and sentiment diversity.

Hypotheses and Research Question

Although there is no compelling proof of decreasing topic diversity over time and although evidence about sentiment diversity is even entirely absent, we postulate two simple longitudinal hypotheses that follow from the literature above and that will guide our analyses in the next sections:

H_1: Topic diversity decreased during the period 2000–2019.

H_2: Sentiment diversity decreased during the period 2000–2019.

In addition, we specifically investigate the possible effects of differences between the four countries in terms of the political and media environment. Most notably, the country sample consists of two different media systems, Liberal in the UK, and Democratic-Corporatist in Denmark, Norway and the Netherlands (Hallin and Mancini 2004). Additionally, the UK has a majoritarian electoral system, while the other countries have a system of proportional representation (Farrell 2011). Also journalistic cultures and role perceptions deviate across the countries, though differences are not strongly pronounced (Hanitzsch et al. 2019). And while the "newspaper crisis" has been found in both Liberal (UK: Lewis, Williams, and Franklin (2008); US: Curran (2010)) and Democratic-Corporatist (Germany: Brüggemann, Esser, and Humprecht (2012); Switzerland: Vogler, Udris, and Eisenegger (2020)) media systems, there might be substantial differences between the UK and the other countries. With its liberal media system and majoritarian electoral system, the highest level of market pressures might be anticipated in the UK. We explore whether this also yields more substantial shifts in diversity than in the Democratic-Corporatist countries. Additionally, the British tabloid newspapers dominate the UK market, with *the Sun* as the most exemplary and outspoken example of this type of newspaper. We explore in detail the behavior of this newspaper and how it relates to the other UK newspapers. Therefore, we formulate an additional research question:

> **RQ1**: What differences are there in the development of diversity between newspapers in the Liberal UK media system and the Democratic-Corporatist systems in the Netherlands, Denmark and Norway?

Data

Sample Selection

To test our hypotheses, we use 6 million newspaper articles, representing 12 newspapers from 4 countries (3 newspapers per country). From each country, a left-leaning broadsheet, right-leaning broadsheet and tabloid/popular newspaper are selected (De Vreese et al. 2016). The selection of newspapers used for each of the four countries under investigation is shown in Table 1. We use all articles from these newspapers, without further sampling.

Pre-processing

All newspaper articles are parsed using Natural Language Processing through the R package UDPipe (Straka and Straková 2017), in combination with version 2.3 of the Dutch Alpino, Danish DDT, English EWT and Norwegian Bokmål Universal Dependencies Models (Nivre et al. 2018). The relevant output of this procedure is a corpus where the

Table 1. Newspaper sample.

Denmark	Netherlands	Norway	UK
Politiken[a]	*de Volkskrant*[a]	*Dagbladet*[d]	*The Guardian*[a]
Jyllands-Posten[b]	*NRC Handelsblad*[b]	*Aftenposten*[b]	*The Daily Telegraph*[b]
Ekstra Bladet[c]	*de Telegraaf*[c]	*VG*[c]	*The Sun*[c]

Note: [a]Left-leaning; [b]Right-leaning; [c]Tabloid/popular; [d]Dagbladet is a left-leaning tabloid rather than a broadsheet.

original (inflected) words have been reduced to their dictionary lemmas. The advantage of using lemmas is that it reduces the number of unique words in the corpus, without losing the substantive meaning of words. In addition, UDPipe produces Universal Part-Of-Speech (UPOS) tags, which identify the grammatical function of a word. The combination of lemma_UPOS pairs is used to train a simple Naive Bayes classification model to remove any articles from the corpora that are not relevant for the current research. These are articles about sports and cultural events, weather forecasts, etc.[1] Such articles do not provide a relevant indication of content diversity from a political/democratic perspective. A full description of the irrelevant article coding procedure and its results can be found in De Vries (2022).

Methods

We use individual content units (articles) as the base unit of analysis and analyze the diversity between those individual units. By aggregating the scores, a measure is constructed, indicating the diversity between newspapers within a country on a given day. We choose to aggregate the article-level comparisons to provide a general overview of the development of content diversity in different countries, over a period of (almost) two decades.[2]

Content Diversity

There are many possible dimensions of content diversity to analyze (see Joris et al. (2020) for an overview). However, because of the large amount of data, and the automated analysis methods used to analyze that data, we do not investigate highly specific dimensions of diversity, such as issue-specific framing. Rather, we focus on word usage, both overall (to what extent do two articles use the same words), and on the sentiment of the words used. The former should capture the general topic of the article, while the latter should capture the sentiment context in which the topic is presented.

To construct a content diversity measure, we compare all articles per day and country to each other, using the lemmas those articles consist of. The ways in which such a measure can be constructed vary; Boumans (2016) uses cosine similarity in combination with tf-idf feature weighting to determine the extent to which newspaper articles are based on external materials, while Vogler, Udris, and Eisenegger (2020) use Jaccard similarity in combination with tri-grams to determine the amount of content sharing between newspapers. The differing approaches in these two studies are in line with their respective goals. Boumans (2016) is focused on newspaper articles that are based on or resemble external materials, while Vogler, Udris, and Eisenegger (2020) are searching for near-duplicate articles. In the former case, the prime goal is to detect (possibly partial) overlap between the article and external material, while in the latter the emphasis is on detecting duplicates. This difference, between overlap and duplicates is also relevant in the context of content and topic diversity.

In their very basis, automated methods for analyzing text are all based on linguistic measures, usually word counts. But through the choice of similarity measure and feature selection/weighting the substantive meaning and interpretation of such measures differs. Jaccard similarity only considers the presence or absence of specific words in a pair of texts, while cosine similarity takes into account the relative frequency of those words. Similarly, on the one hand, tri-grams (groups of three words) are dependent on word

order, and as a result much more strict than regular word frequencies when it comes to measuring similarity. On the other hand, tf-idf weighting as applied by Boumans (2016) increases the weight of more informative words (Jones 1972). Tf-idf weighting is based on the assumption that words occurring in many different articles are less informative than words occurring in only a few articles. Thus, using cosine similarity with idf-weighted word frequencies provides a substantially different measure of similarity than using Jaccard similarity with tri-grams. The former is oriented towards finding overlap, while the latter is oriented towards finding (near-)duplicates (i.e., the exact same words in the exact same order). Because we are interested in the extent to which news articles cover the same events, and not whether one article is a literal copy of another, we use cosine similarity in combination with tf-idf weighting to measure topic diversity. We also compare the sentiment between article pairs that are about the same topic to account for possible differences in the sentiment with which a topic is discussed. As our measure of topic diversity is in its basis a similarity measure, diversity is considered to be the opposite of similarity for the remainder of this paper.

Metrics

To construct the topic diversity measure, we use the functions provided by the R package Quanteda (Benoit et al. 2018). First, the raw articles, consisting of lemmas, are converted into a document-feature matrix, which represents the articles as vectors of word/feature frequencies. This is done by day and country. The feature frequencies are weighted using the idf weighting scheme, and a cosine similarity matrix is constructed. This matrix contains cosine similarity values for every possible article pair on the given day, except the comparisons of articles with themselves. We discard comparisons of articles with other articles from the same newspaper, since we are interested in diversity between rather than within newspapers—external rather than internal diversity. What remains are the similarity values for each article when compared to all articles published in the other two newspapers on the same day. These values are aggregated, so that each article in newspaper A gets a mean and maximum similarity with the articles in newspaper B (or C). These maximum values, indicating the most similar article pairs between newspapers, are used as our indicator for topic diversity.

Although the comparisons between individual articles are by definition symmetric (article A is as much like article B as article B is like article A), this symmetry disappears when using aggregate measures. In our specific application, using the most similar article pairs as indicator of diversity, it is possible that article A might be most like article B, but article B is more like article C than article A. An example of this is provided in Table 2, containing fictional cosine similarity scores between three articles from newspaper A and three articles from newspaper B. This table shows that articles A3 and B3 (.8) and A2 and B1 (.9) are most alike. However, article A1 is most like B1 (.6), while article B2 is most like A1 (.4). On average, then, the similarity of newspaper A with newspaper B is $\frac{.6 + .9 + .8}{3} = .77$ while the similarity of newspaper B with newspaper A is $\frac{.4 + .9 + .8}{3} = .70$. Because of this asymmetry, newspaper comparisons for all countries are made both ways (so A–B, A–C, B–C as well as B–A, C–A, C–B).

Table 2. Similarity matrix example.

	Article B1	Article B2	Article B3
Article A1	.6	**.4**	.1
Article A2	**.9**	.1	.2
Article A3	.1	.3	**.8**

Note: Row maximum is indicated in *italic*, column maximum is indicated in **bold**.

To construct our final topic diversity measure, we use the inverted scores of the most similar article pairs.[3] Inversion of the scores is done to conform to the theoretical concept of diversity, with higher values indicating higher diversity rather than similarity. The article pairs are then split into two categories based on their diversity scores, one with pairs that are about the same topic, and one with pairs that are not. The cutoff value for this split is based on the manual validation results (see below). Then H_1 is tested by evaluating the trend in the weighted percentage[4] of article pairs that are about the same topic.

To test H_2, an indicator for the difference in sentiment between the newspapers is constructed by comparing the sentiment of article pairs that are about the same topic (see above). Article-level sentiment scores are generated using the method described in De Vries (2022). This method consists of a computer-generated dictionary in each of the four languages, used to classify trinary (negative, neutral, positive) sentiment in each sentence of each news article. The sentence-level scores are aggregated and weighted on length to construct a sentiment score at the article level. A sentiment diversity measure is constructed by (1) subtracting the sentiment of article pairs about the same topic from each other, (2) taking the absolute value, and (3) dividing those by 2 (as the original sentiment scores range from +1 to −1). As such, we evaluate the difference in the general sentiment context used to describe a topic. Conceptually, the measure indicates the extent to which articles that already talk about the same topic differ from each other in their general presentation (in terms of positivity/negativity) of that topic.[5] While a similar approach could have adopted for other content features (e.g., presence of actors, or even more substantial frames), we consider general sentiment as a generic and widely studied content feature of media coverage that drives media effects on e.g., public opinion (Vliegenthart et al. 2021). Even though the source and target of the sentiment are unknown, the fact that a topic is discussed within a specific sentiment context is a relevant content feature of media coverage, in particular when we consider external diversity. The degree of congruence in coverage across outlets provides information about similarity, and thus of the level of external diversity as we conceptualize this in our paper. When the sentiment differs between two articles about the same topic, this is an indicator of the diversity of the context within which information on this topic is provided. More specifically, different words imply that a set of articles, even though they have the same topical focus, say something different, and thus add different information and interpretation to the news supply.

Validation

To validate the automatically generated topic diversity metric, we use a small-scale manually coded dataset consisting of 200 Norwegian and 200 UK article pairs. These articles are coded based on the following question: Do these two newspaper articles cover the same

topic? There are two answer options (yes or no). The random sample of article pairs to manually code is constructed in a stratified way based on the automated coding. The topic diversity scores are binned into 5 groups (0–.2, .2–.4, .4–.6, .6–.8, .8–1), and from each of these groups a sample of 40 article pairs is drawn for a total sample of 200 article pairs for Norway and the UK. From the UK dataset, a random subsample of 20 article pairs is used to test the intercoder reliability between 2 coders (the main author and a student assistant). The result of this test is a Krippendorff's alpha of .88, which is more than sufficient. The student assistant, being proficient in both English and Norwegian, has coded the remaining UK article pairs and all 200 Norwegian article pairs.

The human-coded validation results[6] show a strong correlation with our automated topic diversity measure in both the UK ($r(198) = -.79$, $p < .001$) and Norway ($r(198) = -.73$, $p < .001$).

These correlations are visualized as a box plot (Figure 1), showing that topic diversity for article pairs that are about the same topic is generally below .6, while topic diversity for article pairs that are about different topics is generally above .6. Thus a topic diversity value of below .6 (or a cosine similarity value above .4) is used as a cut-off point to determine which article pairs are about the same topic. Extensive validation of the sentiment analysis method used in this paper is presented in De Vries (2022). The general performance ranges from .61 to .64 (weighted F_1), indicating that in a majority of cases sentiment in sentences is classified correctly as either positive, negative or neutral. In addition, the aggregated nature of the current analyses (from sentence to article) and the near-normal distribution of errors allows a majority of the random errors to cancel each other out at the article level.

In addition to the manual validations described above, we have also conducted a face validity test of the topic diversity measure, by comparing the percentage of articles pairs with a topic diversity below .6 and published on the same day to the percentage when one of the articles in a pair is published up to a week later. The reasoning behind this test is that due to the newspaper news cycle spanning a single day, article pairs should
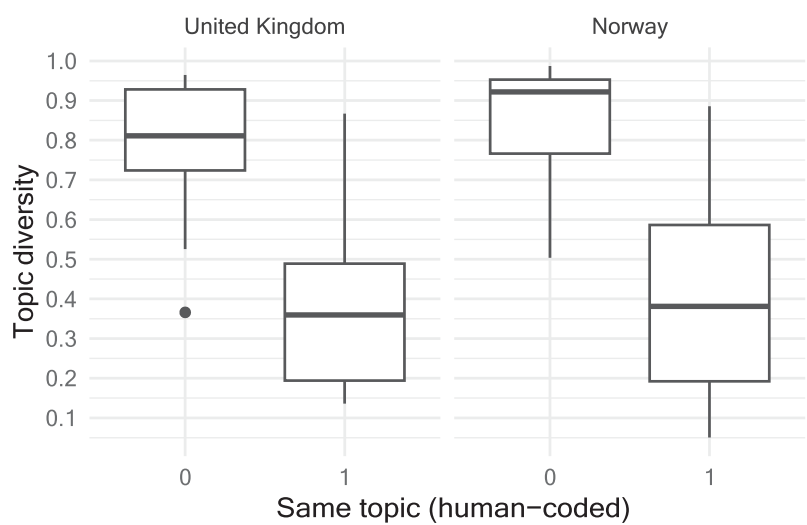

Figure 1. Box plot of topic diversity for articles that are (1) or are not (0) about the same topic.
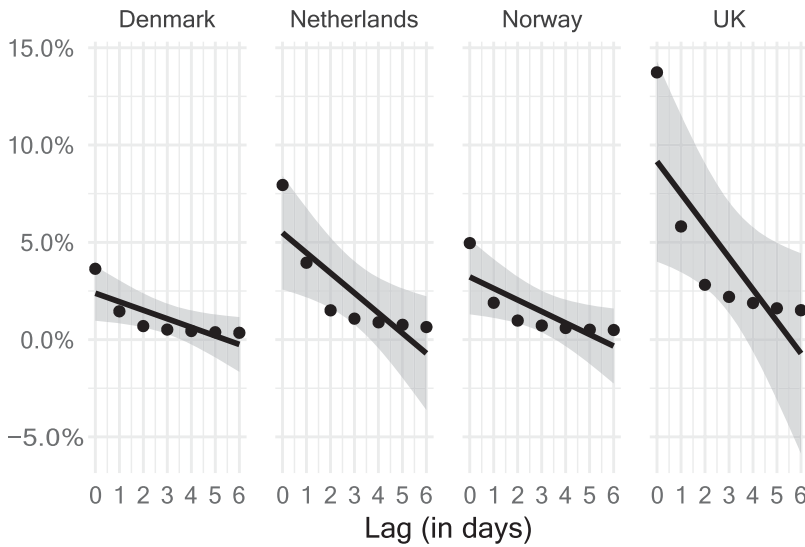
Figure 2. Linear regression showing the effect of publication lag (i.e., article pairs that are published on different days, with a lag from 0 to 6 days) on the percentage of articles with a topic diversity below .6.

be less diverse when they are both published on the same day, than when one is published on a subsequent day. The linear regression results (with standard errors) in Figure 2 test this assumption at the country level. Topic diversity between article pairs published on the same day is indeed lower than when one of the articles is published on a subsequent day, providing additional validity to the topic diversity measure.

Results

In all plots presented below, the black lines show trends, while the grey lines show LOESS-smoothed observations, using a rolling window over 15% of the data. Descriptive statistics of the variables used in the figures and regression models can be found in the appendix (see supplementary online). The results in Figure 3 show the percentage of article pairs that are about the same topic for each newspaper pair (rows), as well as the average per country (columns). On average, there are 116 article pairs per newspaper pair and day, with a standard deviation of 88.

Topic Diversity

As shown in the bottom row of Figure 3, there is for all countries except Denmark a modest decrease in the percentage of article pairs that are about the same topic. In Norway and the Netherlands these trends are quite comparable, while the trend is somewhat more pronounced in the UK and entirely absent in Denmark. In general, no increase in the percentage of article pairs about the same topic is found in any of the countries. Thus the assumption that diversity that decreases over time, as formulated in H_1, is not supported.
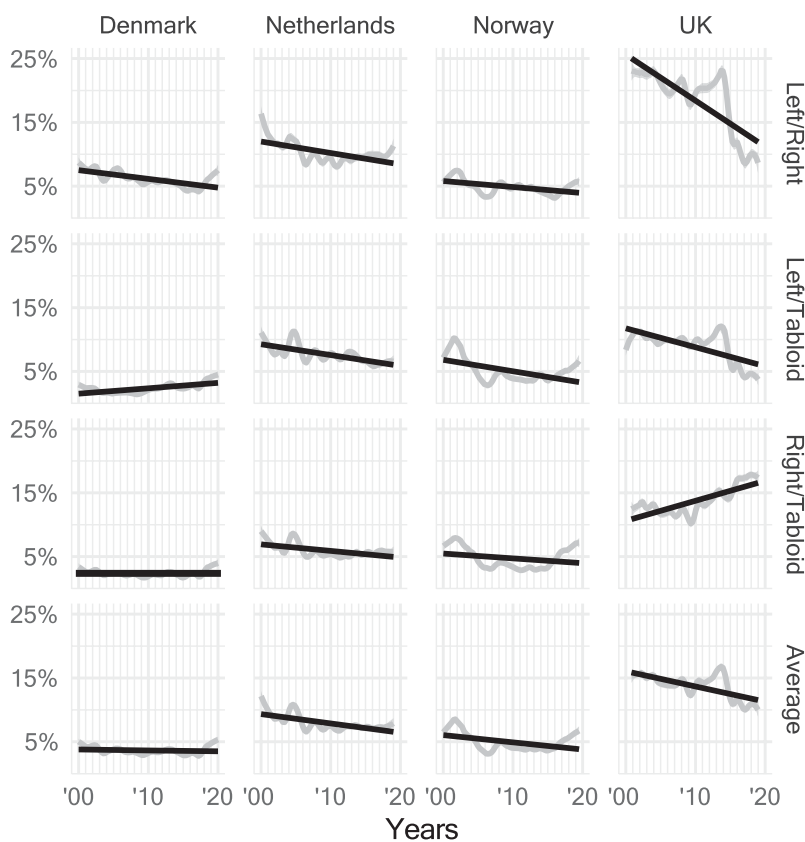
Figure 3. Weighted (by word count) percentage of article pairs with a diversity of .6 or lower, by country and newspaper pair.

A rejection of H_1 is also confirmed by the regression results presented in Table 3. A lagged (by one day) dependent variable is included in these models as a control variable to account for external factors that influence the amount of topic diversity on consecutive days. The standardized regression coefficients show that the effect of time on the percentage of article pairs that share the same topic is in all cases negative and highly significant. The strength of the effect is also comparable between the different countries, except for Denmark, where it is an order of magnitude smaller. Based on the coefficients it is also clear that generally speaking the percentage of article pairs with the same topic is highest between left- and right-wing newspapers and decreases significantly between either left- or right-wing broadsheets and the tabloid newspaper. A notable exception here is Norway, where there is not much difference between the newspaper pairs at all. With an R^2 between .40 (UK) and .09 (Norway) these regression models differ substantially in the amount of variance they explain. In general, a low R^2 is to be expected, as many possible causes for a temporary increase or decrease in topic diversity are not included in these models. However, based on the amount of explained variance, it seems like diversity between newspapers is explained much more by the included predictors in the UK than it is in Norway. Also, the high coefficient of the lagged dependent variable indicates there is more temporal invariance in the amount of article pairs that are

Table 3. Regression results: percentage of article pairs about the same topic.

	Denmark	Netherlands	Norway	UK
Percentage of article pairs (lagged)	.148***	.110***	.250***	.343***
	(.007)	(.008)	(.007)	(.007)
Time (in years)	−.018***	−.146***	−.114***	−.113***
	(.006)	(.007)	(.007)	(.006)
Left-wing/tabloid	−.821***	−.469***	.037**	−.855***
	(.016)	(.017)	(.017)	(.017)
Right-wing/tabloid	−.820***	−.772***	−.021	−.430***
	(.016)	(.018)	(.016)	(.016)
Constant	.547***	.411***	−.005	.445***
	(.011)	(.012)	(.012)	(.012)
Observations	21,686	17,291	20,097	16,688
R^2	.224	.166	.085	.396
Adjusted R^2	.224	.166	.085	.396
Residual std. error	.881	.913	.957	.777
F statistic	1564.952***	861.050***	468.263***	2734.278***

*$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

about the same topic in the UK than in any of the other countries. In Norway, the included variables do not add to the explanation for the amount of diversity at all, except for time.

Sentiment Diversity

In contrast to the findings relating to topic diversity, there are no clear trends visible in the bottom row of Figure 4. This figure shows for each country the average normalized level of sentiment diversity between articles that are about the same topic. There do not appear to be any substantial differences between the countries, which is supported by the standardized regression results in Table 4. As the explained variance of the Norwegian regression model is exactly 0, we disregard this model entirely with the comment (like with topic diversity) that the included variables do not predict sentiment diversity at all. Regarding the other countries, the most notable coefficients in these regressions are those indicating the newspaper pairs. All of these are comparable in size and significance between the countries, indicating that the left- and right-wing newspapers differ substantially more from the tabloid newspaper than from each other. In general, the

Table 4. Regression results: sentiment diversity.

|	Denmark	Netherlands	Norway	UK
Sentiment diversity (lagged)	.006	.008	.001	−.009
	(.008)	(.008)	(.008)	(.008)
Time (in years)	.012	−.013*	−.010	.121***
	(.008)	(.008)	(.008)	(.008)
Left-wing/tabloid	.139***	.190***	.001	.244***
	(.019)	(.019)	(.019)	(.019)
Right-wing/tabloid	.171***	.207***	.005	.232***
	(.020)	(.019)	(.018)	(.019)
Constant	−.091***	−.127***	−.004	−.163***
	(.012)	(.013)	(.013)	(.014)
Observations	14,328	16,566	17,050	16,477
$R2$	.007	.009	.000	.027
Adjusted $R2$	.007	.009	−.000	.027
Residual std. error	.962	.994	.989	.978
F statistic	24.851***	39.193***	.498	113.992***

*$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.
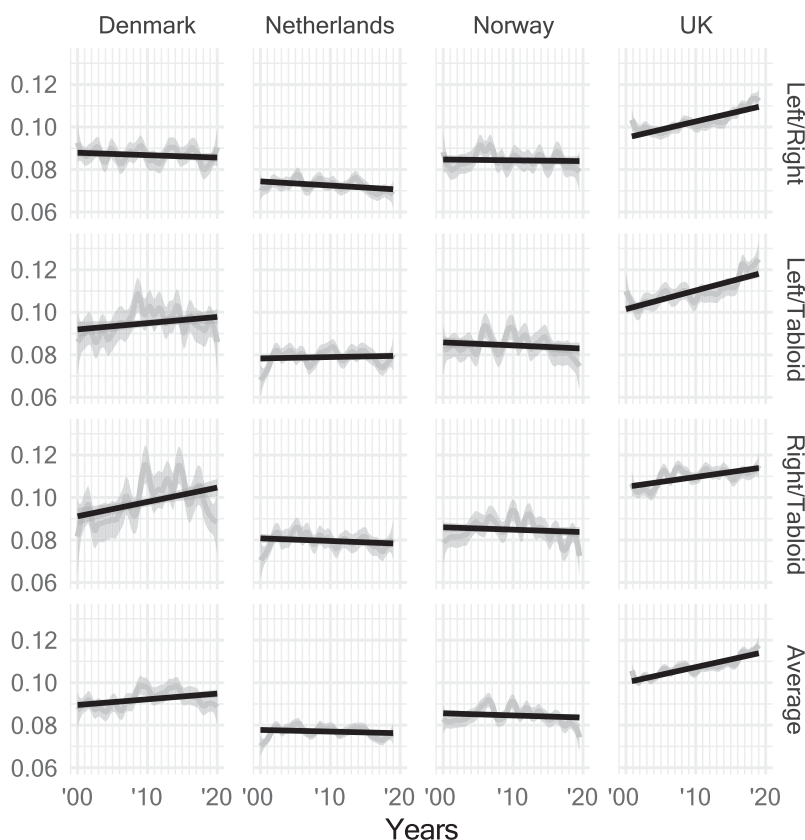
Figure 4. Sentiment diversity of article pairs with a diversity of .6 or lower, by country and newspaper pair.

absence of clear trends is also reflected in the insignificant effect of time and more generally in the negligible amount of explained variance in each of the models. The one exception is the UK, where time does have a significant positive effect on sentiment diversity. In combination with Figure 4 these results lead to a rejection of H_2 as there is no clear decrease of sentiment diversity over time in any of the countries.

Cross-national Comparison

Considering the results presented above, and notwithstanding its different media and political system, the UK does not stand out at country level when it comes to topic diversity (bottom row in Figure 3), and the downward trend is remarkably similar to that in other countries. However, a systematic newspaper by newspaper comparison in the various countries as presented in Figure 3 does reveal a different trend in the United Kingdom: topic diversity between the right-leaning (*The Daily Telegraph*) and tabloid (*The Sun*) newspaper decreases substantially, while it increases substantially between these newspapers and the left-leaning Guardian. Additionally the results in Figure 4, while not indicating substantial differences between the UK newspapers, do show a modest trend towards increasing sentiment diversity. This indicates that the UK

newspapers increasingly differ in the sentiment with which they cover topics, even when they cover the same topics. Both results seem to be indicative for a kind of topic and sentiment polarization or segmentation, which is not found in any of the other countries. In answer to RQ_1 there indeed seems to be a polarizing effect of the political and media system in the UK, even though it is not visible in the country level analyses.

Conclusion

Our study scrutinizes an often-made assumption about journalistic content—that diversity has decreased due to developments such as increasing competition and economic pressures. We put this assumption to a rigid, cross-national empirical test and despite plausible and compelling claims, we find little evidence for a decreasing trend in diversity. If anything, we find evidence for *increasing* topic diversity. The quickly changing media environment might thus not have had the anticipated effect. Potentially, we can attribute the lack of a decreasing diversity trend to the changing role of traditional media, and in particular newspapers. Mellado and colleagues (2017) demonstrate that journalistic role perceptions have become increasingly hybrid. Their study provides compelling evidence about the multilayered hybridization of journalistic cultures at the performative level. Professional roles are varied as well as fluid and dynamic. Compared to several decades ago, the importance of "bringing the news (first)" has substantially decreased for traditional media. Online and social media have taken over the role of being the first ones to bring news to large audiences. Printed newspapers will only in a minority of instances be the source that brings news and events first. They have moved in the direction of providing interpretation, analysis and opinions instead (Esser and Umbricht 2014; Soontjens, 2019; e.g., Strömbäck and Aalberg 2008). This new, more interpretative, role inherently goes hand in hand with a diversification of content, newspapers developing a more distinct profile, that might have canceled out the pressure to less diverse content. It also emphasizes the continuing relevance of newspapers, as they offer something that other news sources do not. In this way, diversity between newspapers continues to affect the news diversity of the entire media landscape. Future research could try to link the results of our study with longitudinal and cross-national data on (changing) role perceptions, for example from the Worlds of Journalism project (Hanitzsch et al. 2019), or to a more detailed analysis of editorial strategies of individual media outlets.

Our findings are indicative of increased instead of decreased diversity, but they do not provide definite answers. First, it might be that decreases in diversity have taken place before the period we scrutinized. After all, changes in the media landscapes and financial pressures originate from well before the end of the previous century (Lewis, Williams, and Franklin 2008). Practical constraints, most notably the absence of digital archives for pre-2000 content for a substantial part of our sources, refrain us from establishing whether this is indeed the case.

Second, our measure of diversity is not comprehensive—we focus solely on external diversity and we look only at diversity in terms of topics and sentiment. We do not look at actors, for instance, or at framing. It might well be that external topic diversity is high according to our measure but that internal diversity is low; the readers of a given newspaper might be confronted with low internal diversity although they could find more diverse news by looking also at other outlets. Thus, our measure only partially

grasps what actual news consumers are confronted with, as only part of them read several newspapers. This limitation, next to the lack of a comparison with the actual diversity in preferences and opinions in society (through e.g., survey research), thwarts the opportunity to make a comprehensive assessment of our findings and compare them against ideal types such as reflective and open diversity (Van Cuilenburg 1999).

Third, while our study is cross-national and we engaged in some tentative comparative analyses, we have not been able to truly capitalize on this in terms of providing a systematic comparative account. We would need more countries to statistically test the impact of country level features. Nonetheless, we have found some tantalizing evidence that the increased competition in the Liberal media system of the UK (Hallin and Mancini 2004) does not lead to relatively less diversity over time than in the Democratic-Corporatist media systems in Denmark, The Netherlands and Norway. Rather, it seems to lead to a more polarized form of diversity, with one newspaper being substantially different from the other two.

Finally, the interpretability of our results is limited because our news diversity measures disregards word order and syntax, and are measured at the article level rather than for example sentence level. The sentiment diversity measure therefore indicates the difference in general tone when the same topic is discussed, but does not account for viewpoint diversity within articles, nor the sentiment associated with specific viewpoints. For topic diversity, the loss of word order and syntax opens up the hypothetical possibility that two articles either use very similar words to describe very different topics or use very dissimilar words to describe very similar topics. However, such cases do not seem likely, as our topic diversity measure (a combination of cosine similarity with tf-idf feature weighting) is shown to be reliable and valid. It strongly correlates with human classifications, works well with languages other than English, and is not overly complex or computationally expensive.

In addition, our study is the first to show trends of increasing news diversity while combining a multiple country comparison with a nearly two-decade time span. One thing that stands out in particular is that the findings are largely similar across countries. In that sense, our paper offers a robust assessment of general patterns that hold across contexts. Patterns, as shown here, that turned out to be quite different from our theoretical expectations.

Notes

1. See the appendix (supplementary online) for a full overview.
2. The supplementary materials (scripts and data) to replicate our findings are available at https://osf.io/3hdk4/. Note that due to copyright restrictions, we are unable to share the raw data containing the full article texts, with the exception of the human validation datasets.
3. i.e., the articles in A that are most like the articles in B or C, and vice versa.
4. by the average article length in number of words.
5. Variations in the sentiment related to different viewpoints within a topic are not considered, due to measurement at the article level.
6. Short examples of articles at various levels of topic diversity can be found in the appendix, while more extensive examples can be found in the validation dataset that is part of the supplementary materials at https://osf.io/3hdk4/.

Disclosure Statement

ORCID

Erik de Vries 🔵 http://orcid.org/0000-0002-4961-6047
Rens Vliegenthart 🔵 http://orcid.org/0000-0003-2401-2914
Stefaan Walgrave 🔵 http://orcid.org/0000-0003-1814-1015

References

Amsalem, E., Y. Fogel-Dror, S. R. Shenhav, and T. Sheafer. 2020. "Fine-Grained Analysis of Diversity Levels in the News." *Communication Methods and Measures* 14 (4): 266–284. doi:10.1080/19312458.2020.1825659.

Baker, C. E. 2007. *Media Concentration and Democracy: Why Ownership Matters*. Cambridge: Cambridge University Press.

Beckers, K., A. Masini, J. Sevenans, M. van der Burg, J. De Smedt, H. Van den Bulck, and S. Walgrave. 2019. "Are Newspapers' News Stories Becoming More Alike? Media Content Diversity in Belgium, 1983." *Journalism* 20 (12): 1665–1683. doi:10.1177/1464884917706860.

Benoit, K., K. Watanabe, H. Wang, P. Nulty, A. Obeng, S. Müller, and A. Matsuo. 2018. "Quanteda: An R Package for the Quantitative Analysis of Textual Data." *Journal of Open Source Software* 3 (30): 774. doi:10.21105/joss.00774.

Boczkowski, P. J., and M. de Santos. 2007. "When More Media Equals Less News: Patterns of Content Homogenization in Argentina's Leading Print and Online Newspapers." *Political Communication* 24 (2): 167–180. doi:10.1080/10584600701313025.

Boukes, M., B. Van de Velde, T. Araujo, and R. Vliegenthart. 2020. "What's the Tone? Easy Doesn't Do It: Analyzing Performance and Agreement Between Off-the-Shelf Sentiment Analysis Tools." *Communication Methods and Measures* 14 (2): 83–104. doi:10.1080/19312458.2019.1671966.

Boumans, J. 2016. *Outsourcing the News? An Empirical Assessment of the Role of Sources and News Agencies in the Contemporary News Landscape*. Amsterdam: University of Amsterdam.

Boumans, J., D. Trilling, R. Vliegenthart, and H. Boomgaarden. 2018. "The Agency Makes the (Online) News World Go Round: The Impact of News Agency Content on Print and Online News." *International Journal of Communication* 12: 22.

Brüggemann, M., F. Esser, and E. Humprecht. 2012. "The Strategic Repertoire of Publishers in the Media Crisis." *Journalism Studies* 13 (5-6): 742–752. doi:10.1080/1461670X.2012.664336.

Curran, J. 2010. "The Future of Journalism." *Journalism Studies* 11 (4): 464–476. doi:10.1080/14616701003722444.

Dailey, L., L. Demo, and M. Spillman. 2005. "The Convergence Continuum: A Model for Studying Collaboration Between Media Newsrooms." *Atlantic Journal of Communication* 13 (3): 150–168. doi:10.1207/s15456889ajc1303_2.

Damstra, A., and K. De Swert. 2020. "The Making of Economic News: Dutch Economic Journalists Contextualizing Their Work." *Journalism*. doi:10.1177/1464884919897161.

Day, A. G., and G. Golan. 2005. "Source and Content Diversity in Op-Ed Pages: Assessing Editorial Strategies in The New York Times and the Washington Post." *Journalism Studies* 6 (1): 61–71. doi:10.1080/1461670052000328212.

De Vreese, C. H., F. Esser, D. N. Hopmann, F. Esser, and D. N. Hopmann. 2016. *Comparing Political Journalism*. Routledge. doi:10.4324/9781315622286

De Vries, E. 2022. "The Sentiment is in the Details. A Language-Agnostic Approach to Sentence-Level Sentiment Analysis in News Media." *Computational Communication Research* 4 (2): 424–447. doi:10.5117/CCR2022.2.003.VRIE.

Entman, R. M. 2003. "Cascading Activation: Contesting the White House's Frame After 9/11." *Political Communication* 20 (4): 415–432. doi:10.1080/10584600390244176.

Esser, F., and A. Umbricht. 2014. "The Evolution of Objective and Interpretative Journalism in the Western Press: Comparing Six News Systems Since the 1960s." *Journalism & Mass Communication Quarterly* 91 (2): 229–249. doi:10.1177/1077699014527459.

Farrell, D. M. 2011. *Electoral Systems: A Comparative Introduction*. London: Bloomsbury Publishing.

Gandy, O. H. 1980. "Information in Health: Subsidised News." *Media, Culture & Society* 2 (2): 103–115. doi:10.1177/016344378000200201.

Hallin, D. C., and P. Mancini. 2004. "Comparing Media Systems: Three Models of Media and Politics." In *Cambridge Core*. Cambridge University Press. doi:10.1017/CBO9780511790867.

Hanitzsch, T., F. Hanusch, J. Ramaprasad, and A. S. de Beer, eds. 2019. *Worlds of Journalism: Journalistic Cultures Around the Globe*. New York: Columbia University Press.

Hanusch, F. 2015. "Transformative Times: Australian Journalists' Perceptions of Changes in Their Work." *Media International Australia* 155 (1): 38–53. doi:10.1177/1329878X1515500106.

Hendrickx, J., and H. Ranaivoson. 2019. "Why and How Higher Media Concentration Equals Lower News Diversity The Mediahuis Case." *Journalism*. doi:10.1177/1464884919894138.

Hotelling, H. 1929. "Stability in Competition." *The Economic Journal* 39 (153): 41–57. doi:10.2307/2224214.

Jones, K. S. 1972. "A Statistical Interpretation of Term Specificity and its Application in Retrieval." *Journal of Documentation* 28: 11–21.

Joris, G., F. D. Grove, K. V. Damme, and L. D. Marez. 2020. "News Diversity Reconsidered: A Systematic Literature Review Unraveling the Diversity in Conceptualizations." *Journalism Studies* 21 (13): 1893–1912. doi:10.1080/1461670X.2020.1797527.

Jyrkiäinen, J., and A. Heinonen. 2012. "Finnish Journalists: The Quest for Quality Amidst New Pressures." In *The Global Journalist in the 21st Century*, edited by David H. Weaver and Lars Willnat, 171–186. New York: Routledge.

Lacy, S., and T. F. Simon. 1993. *The Economics and Regulation of United States Newspapers*. New York: Ablex.

Langer, A. I., and J. B. Gruber. 2021. "Political Agenda Setting in the Hybrid Media System: Why Legacy Media Still Matter a Great Deal." *The International Journal of Press/Politics* 26 (2): 313–340. doi:10.1177/1940161220925023.

Lee, J. K. 2007. "The Effect of the Internet on Homogeneity of the Media Agenda: A Test of the Fragmentation Thesis." *Journalism & Mass Communication Quarterly* 84 (4): 745–760. doi:10.1177/107769900708400406.

Lewis, J., A. Williams, and B. Franklin. 2008. "Four Rumours and an Explanation." *Journalism Practice* 2 (1): 27–45. doi:10.1080/17512780701768493.

Masini, A., and P. Van Aelst. 2017. "Actor Diversity and Viewpoint Diversity: Two of a Kind?" *Communications* 42 (2): 107–126. doi:10.1515/commun-2017-0017.

Mellado, C., L. Hellmueller, M. Márquez-Ramírez, M. L. Humanes, C. Sparks, A. Stepinska, Svetlana Pasti, Anna-Maria Schielicke, Edson Tandoc, and H. Wang. 2017. "The Hybridization of Journalistic Cultures: A Comparative Study of Journalistic Role Performance." *Journal of Communication* 67 (6): 944–967. doi:10.1111/jcom.12339.

Napoli, P. M. 1999. "Deconstructing the Diversity Principle." *Journal of Communication* 49 (4): 7–34. doi:10.1111/j.1460-2466.1999.tb02815.x.

Newman, N., R. Fletcher, A. Schulz, S. Andi, C. T. Robertson, and R. K. Nielsen. 2021. "Reuters Institute Digital News Report 2021." Reuters Institute for the Study of Journalism.

Nivre, J., M. Abrams, Ž Agić, L. Ahrenberg, L. Antonsen, K. Aplonova, M. J. Aranzabe, et al. 2018. "Universal Dependencies 2.3."

Paulussen, S. 2012. "Technology and the Transformation of News Work: Are Labor Conditions in (Online) Journalism Changing?" In *The Handbook of Global Online Journalism*, edited by E. Siapera and A. Veglis, 192–208. Wiley-Blackwell. doi:10.1002/9781118313978.ch11.

Picard, R. G. 2014. "Twilight or New Dawn of Journalism?" *Journalism Practice* 8 (5): 488–498. doi:10.1080/17512786.2014.905338.

Raeymaeckers, K., S. Paulussen, and J. De Keyser. 2012. "A Survey of Professional Journalists in Flanders (Belgium)." In *The Global Journalist in the 21st Century*, edited by David H. Weaver and Lars Willnat, 141–154. New York: Routledge.

Rodgers, S., E. Thorson, and M. Antecol. 2000. "'Reality' in the St. Louis Post-Dispatch." *Newspaper Research Journal* 21 (3): 51–68. doi:10.1177/073953290002100305.

Roessler, P. 2007. "Media Content Diversity: Conceptual Issues and Future Directions for Communication Research." *Annals of the International Communication Association* 31 (1): 464–520. doi:10.1080/23808985.2007.11679073.

Schudson, M. 2011. *The Sociology of News*. 2nd ed. New York: WW Norton.

Sjøvaag, H. 2016. "Media Diversity and the Global Superplayers: Operationalising Pluralism for a Digital Media Market." *Journal of Media Business Studies* 13 (3): 170–186. doi:10.1080/16522354.2016.1210435.

Soontjens, K. 2019. "The Rise of Interpretive Journalism." *Journalism Studies* 20 (7): 952–971. doi:10.1080/1461670X.2018.1467783.

Soroka, S., and S. McAdams. 2015. "News, Politics, and Negativity." *Political Communication* 32 (1): 1–22. doi:10.1080/10584609.2014.881942.

Straka, M., and J. Straková. 2017. "Tokenizing, POS Tagging, Lemmatizing and Parsing UD 2.0 with UDPipe." In *Proceedings of the CoNLL 2017 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies*, 88–99.

Strömbäck, J., and T. Aalberg. 2008. "Election News Coverage in Democratic Corporatist Countries: A Comparative Study of Sweden and Norway." *Scandinavian Political Studies* 31 (1): 91–106. doi:10.1111/j.1467-9477.2008.00197.x.

Van Aelst, P., J. Strömbäck, T. Aalberg, F. Esser, C. de Vreese, J. Matthes, D. Hopmann, et al. 2017. "Political Communication in a High-Choice Media Environment: A Challenge for Democracy?" *Annals of the International Communication Association* 41 (1): 3–27. doi:10.1080/23808985.2017.1288551.

Van Aelst, P., and S. Walgrave. 2016. "Information and Arena: The Dual Function of the News Media for Political Elites." *Journal of Communication* 66 (3): 496–518. doi:10.1111/jcom.12229.

Van Cuilenburg, J. 1999. "On Competition, Access and Diversity in Media, Old and New: Some Remarks for Communications Policy in the Information Age." *New Media & Society* 1 (2): 183–207. doi:10.1177/14614449922225555.

Vliegenthart, R., A. Damstra, M. Boukes, and J. Jonkman. 2021. *Economic News: Antecedents and Effects*. Cambridge: Cambridge University Press.

Vogler, D., L. Udris, and M. Eisenegger. 2020. "Measuring Media Content Concentration at a Large Scale Using Automated Text Comparisons." *Journalism Studies*, 1–20. doi:10.1080/1461670X.2020.1761865.

Walgrave, S., A. E. Boydstun, R. Vliegenthart, and A. Hardy. 2017. "The Nonlinear Effect of Information on Political Attention: Media Storms and US Congressional Hearings." *Political Communication* 34 (4): 548–570. doi:10.1080/10584609.2017.1289288.