



UvA-DARE (Digital Academic Repository)

The agency makes the (online) news world go round: The impact of news agency content on print and online news

Boumans, J.; Trilling, D.; Vliegenthart, R.; Boomgaarden, H.

Publication date

2018

Document Version

Final published version

Published in

International Journal of Communication : IJoC

License

CC BY-NC-ND

[Link to publication](#)

Citation for published version (APA):

Boumans, J., Trilling, D., Vliegenthart, R., & Boomgaarden, H. (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 : IJoC*, 12, 1768-1789.
<https://ijoc.org/index.php/ijoc/article/view/7109>

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.

The Agency Makes the (Online) News World Go Round: The Impact of News Agency Content on Print and Online News

JELLE BOUMANS¹
DAMIAN TRILLING
RENS VLIAGENTHART
University of Amsterdam, Netherlands

HAJO BOOMGAARDEN
University of Vienna, Austria

While it is generally acknowledged that news agencies play a pivotal role in the current news landscape, empirical insights into the extent of news media's reliance on agency copy are scarce. This study applies an innovative automated approach to trace agency copy for an entire year ($n = 119,452$) in the major print and online news media articles ($n = 247,161$) of the Dutch news landscape. Results suggest that particularly online news is highly dependent on agency content, with the agency being responsible for up to 75% of the online news articles. Furthermore, a large part of the online news consists of verbatim agency copy, involving little or no editing. The results provide a strong rationale to place news agencies high on the agenda of news production scholars. Moreover, the demonstrated agency domination of online news is alarming in the context of news diversity.

Keywords: news production, news agency, intermedia agenda setting, churnalism, transparency, automated content analysis, online, print

News agencies have a long history in the world of the news. Developed as a joint initiative by newspapers in the 1830s and 1840s, the original purpose of the agency was to lower the costs and expand the scope of foreign correspondence. In the past decade, however, their reach has increased

Jelle Boumans: j.w.boumans@uva.nl

Damian Trilling: d.c.trilling@uva.nl

Rens Vliegenthart: r.vliegenthart@uva.nl

Hajo Boomgaarden: hajo.boomgaarden@univie.ac.at

Date submitted: 2017-02-15

¹ This work was carried out on the Dutch national e-infrastructure with the support of SURF Cooperative. For the analysis, we relied on the Infrastructure for Content Analysis of the Amsterdam School of Communication Research.

Copyright © 2018 (Jelle Boumans, Damian Trilling, Rens Vliegenthart, and Hajo Boomgaarden). Licensed under the Creative Commons Attribution Non-commercial No Derivatives (by-nc-nd). Available at <http://ijoc.org>.

dramatically: Today's agencies can be described as "wholesale news providers" (Paterson, 2007), covering everything from politics to sports and from financial to entertainment news. They have a reputation of delivering maximally accurate, maximally fast, and maximally factual information (Boyd-Barrett, 2010; Boyer, 2011), and, according to the editor of the British national agency Press Association, they are "the central heart of the media industry" (Davies, 2008, p. 74). News media increasingly rely on news agencies to cope with the 24/7 news cycle and the pressure to be first with the news (Karlsson, 2011) as well as to reduce their production costs (Boyer, 2011; Forde & Johnston, 2013). This reliance is even stronger for online journalism (Johnston, 2009; Paterson, 2007). Unlike most of the other information that a journalist draws upon to produce a story, information from agencies has an authority so absolute that the news media generally do not question its content or factual accuracy (Davies, 2008; Johnston & Forde, 2011).

The (predominantly qualitative) research on news agencies to date has provided a rich account of the historical and current dominant position of news agencies in the news landscape (Boyd-Barrett, 2000; Esperidião, 2011; Horvit, 2006; Paterson, 2011). Although several attempts have been made to quantify the degree to which news media rely on agency copy (Forde & Johnston, 2013; Johnston & Forde, 2011; Lewis, Williams, & Franklin, 2008; Paterson, 2001, 2007), this has seldom been done on a scale that allows generalization across outlets and news sections. We argue that more robust evidence is required to fully understand how agencies affect media agendas. To this end, this study investigates the influence of these "silent partners of the news" (Johnston & Forde, 2011, p. 196) in the Dutch context, providing evidence on a national scale for what is believed to be a global trend.

The automated approach that was developed specifically for this purpose enables us to assess with unprecedented accuracy the impact of the principal national agency on the news. It distinguishes between two levels of the news agency's influence: first, on the media agenda level (i.e., To what extent is the news agenda initiated by agency copy?) and, second, on the content level (i.e., To what extent does news content consist of verbatim agency text?). Given the scarcity of empirical data on these matters, our general research question is descriptive in nature:

RQ1: To what extent are the national news agenda and news content based on agency copy?

In theoretical terms, the first level is embedded in the intermedia agenda-setting framework, which is concerned with how news media affect attention for issues. It is the mechanism that creates a common definition of what news is and is not (Vliegenthart & Walgrave, 2008). The second level is related to the notion of "churnalism"—"a form of journalism that relies on recycling press releases and agency copy and which involves little or no independent reporting or attempt at verification" (Harcup, 2014, p. 53). This study focuses on the use of agency copy; for an assessment of newspapers' reliance on press release material, see Boumans (2017).

The few empirical studies that have traced agency copy in newspaper content found that the reliance of news organizations on agency copy is rarely attributed (Hijmans, Schafraad, Buijs, & d'Haenens, 2011; Johnston, 2009; Lewis, Williams, & Franklin, 2008; Scholten & Ruigrok, 2009). The automated analysis is able to detect attribution (or lack thereof), allowing for a highly accurate

observation of the extent to which Dutch print and online news media perform on this indicator of transparency (Karlsson, 2011).

Reliance on agency content has been studied for television news (Paterson, 2011), print news (Lewis, Williams, & Franklin, 2008), and online news (Johnston, 2009), but comparative research between media platforms is scarce. To our knowledge, only one comparative content analysis has been published; the study demonstrates that online news media rely more on agency copy than newspapers do (Welbers, Van Atteveldt, Kleinnijenhuis, & Ruigrok, 2016). Whereas Welbers and colleagues focused on political news, the current study includes the broad spectrum of news sections.

Apart from the aim to inform the debate on agency reliance with empirical evidence, the study makes a methodological contribution by introducing an innovative automated tool that enables us to trace agenda setting and churnalism patterns in large-scale data sets. Reliance on news agency copy is generally measured by counting the references to the agency in the credit line (Hijmans et al., 2011; Powers & Benson, 2014; Sjøvaag, 2014; Van Leuven, Deprez, & Raeymaeckers, 2014), which is a notoriously inaccurate measurement given news organizations' inconsistent attributing practices (Quandt, 2008; Reich, 2010). The rare study that considers the actual agency copy takes a case study approach (Johnston & Forde, 2011), limiting the generalizability. Notable exceptions aside (Lewis, Williams, Franklin, Thomas, & Mosdell, 2008), the generalizability issue also applies to research on churnalism.

This is not the first attempt to quantify the use of agency copy using computer-assisted text comparison: Paterson (2001, 2007) has demonstrated in a small-scale computer-assisted analysis that online news content shows large overlap with agency copy. Although the approach enabled the researcher to move beyond case studies, the still predominantly manual method proved to be labor-intensive and hence less suitable for large-scale analyses. In contrast, the novel design of the approach presented here allows for scaling up the analysis to a data set unprecedented in terms of size and scope. Furthermore, by applying standardized measures to the agency's impact on the print and online news landscape, we hope to facilitate future (comparative) research.

The Central Role of the Agency in the News Landscape

The dominant role of international news agencies in the construction of global news has been demonstrated in various studies. Although research on national news agencies is somewhat scattered, Boyd-Barrett's (2010) edited volume provides a rich account of the role and future of national news agencies from across the world. Furthermore, ethnographic studies of agency newsrooms provide detailed insights into agency journalists' routines and the agencies' capacity to assemble, produce, and disseminate information (Boyer, 2011; Paterson, 2011). What emerges from this varied body of research on news agencies is a picture of the news agency as a central hub in the world of the news, yet exactly *how* large news media rely on agency copy has seldom been quantified (but see Paterson, 2007; Welbers et al., 2016). Weaver and Bimber (2008) have suggested that the scarcity of systematic and large-scale "input-output" analyses of agency reliance could possibly be explained by a lack of data availability. In the United States, for instance, an agency story is featured only once in the LexisNexis database due to licensing agreements, regardless of whether the same story has been replicated in multiple newspapers.

However, for a short period (2002–7), Google News did retrieve all U.S. newspaper articles that were replications of agency copy, providing a unique opportunity to assess the distribution of agency content. Comparing the retrieved articles of both databases, LexisNexis was found to be missing half or more of the stories appearing in major newspapers because it is “blind to wire stories” (Weaver & Bimber, 2008, p. 515). The study thus demonstrated not only that newspapers routinely use agency copy but that assessing this reliance can be hindered by data availability issues.

Criticism of News Agencies’ Functioning

Although it is widely agreed that agencies generally succeed in fulfilling their pivotal societal task of providing accurate and factual information from across the globe, several concerns about the functioning of news agencies have been raised. Early criticism in particular focused on the hegemonic role of international news agencies (Boyd-Barrett, 1980), signaling an “overwhelming Euro-American dominance of global news flow” (Boyd-Barrett, 2000, p. 12). More recently, concerns are expressed about the agencies’ alleged inability to cover the entire spectrum of potential news events (Davies, 2008; Esperidião, 2011) and the detrimental impact that agencies have on news diversity (Johnston, 2009; McChesney, 2003, 2008; Paterson, 2007, 2011). The key concern here is that “relaying a limited selection of information in a limited set of ways might limit public understanding” (Paterson, 2011, p. 156). It has furthermore been suggested that agency content is relatively more reliant on routine and elite sources (Esperidião, 2011; Manning, 2008), is less critical (Forde & Johnston, 2013; Jackson & Moloney, 2016), and is more polarized in tone (Ellis, 2007).

News Agencies Under Economic Pressure

News agencies—and in particular wire services—are facing challenging times. The agencies have lost their traditional monopoly on “raw” news and find themselves faced with multiple parties that operate on their market (Boyer, 2011). In the digital age, it is increasingly difficult to secure exclusivity of the content: As soon as a medium publishes content online, other media are able to profit for free (Tijdelijke Commissie Innovatie en Toekomst Pers, 2009). In addition, as a consequence of media organizations’ cost orientation, the agencies are under increased pressure to deliver more services at lower prices (Phillips, 2010a) and experience fewer turnovers as media clients unsubscribe from agency services (Wilke, 2010). Agencies worldwide respond to these challenges by cutting personnel and seeking additional revenues through diversifying their services (Boyd-Barrett, 2010; Manning, 2008; Rutten & Slot, 2011). These activities often include photo, video, and mobile media services, but also services that are criticized for crossing the line between news and public relations (Paterson, 2010; Tijdelijke Commissie Innovatie en Toekomst Pers, 2009; Wilke, 2010).

Churnalism

The increased reliance of news organizations on agency copy is often explained by referring to journalists’ increased workload (Boyer, 2011; Johnston & Forde, 2011) and the need to cope with the “news cyclone” (Klinenberg, 2005). According to Davies (2008), many journalists are in fact information brokers who mainly recycle existing content—a practice that is dubbed *churnalism*. Although different

terms have been used to describe this phenomenon (for an overview, see Van Leuven et al., 2014), the critique is consistent: Journalists habitually incorporate subsidized content from agencies, other media, or public relations professionals into their news output (Davis, 2000; Jackson & Moloney, 2016; Lewis, Williams, Franklin, Thomas, et al., 2008; Reich, 2010). Critics are concerned that the reactive and deskbound practice of today's journalism leads to a decline in the diversity of viewpoints in the news, as less-resourced voices are unable to make themselves heard in the "Niagara of PR propaganda" (Moloney, 2006, p. 2) that floods the journalistic system. The growing dominance of agencies in the news production process is feared to further compromise the diversity of voices and perspectives in the news: The limited time to evaluate the overload of information combined with the agency's primary goal of delivering factual, trustworthy information makes news agency journalists particularly susceptible to relying on established sources (Boyer, 2011; Livingston & Bennett, 2003).

Online Versus Off-Line News

Online news is increasingly the primary source of information, particularly for young consumers (e.g., Mitchelstein & Boczkowski, 2010; Trilling & Schoenbach, 2015). Web journalists have different work routines and role conceptions than do print journalists (Paterson, 2011). Technological and market restrictions as well as different resources, tools, and working routines all contribute to online media reporting being generally less meticulous than print media. Often lacking working business models (Humprecht & Buchel, 2013), it is assumed that online media in general rely strongly on the most efficient way of producing news: disseminating agency copy. Some view the Web as an alternative news source, while others claim it is an alternative market for agency copy (Doyle, 2015; Fenton, 2010; Johnston, 2009). According to Paterson (2007), who analyzed the reliance on news agencies in the context of international news, the Internet presents "the dangerous illusion of multiple perspectives which actually emanate from very few sources" (p. 64). A cross-national study on print and online news homogeneity found that the prevalence of news agency copy is up to four times higher in online news than it is in print in all three studied countries (the United States, France, and Denmark; Powers & Benson, 2014). Yet this observation was made on the basis of counting explicit references to the agency in the byline—a method that has proven to be notoriously unreliable (Quandt, 2008). The approach proposed here compares the actual content of the agency and newspaper articles. Guided by the previous findings (see also Welbers et al., 2016), we formulate the following hypotheses:

H1: Print news will have a lower percentage than online news of articles that are initiated by agency copy.

H2: The verbatim replication of news agency content will occur less in print news than in online news.

Attribution to the News Agency

In a climate of declining public trust in journalism, transparency is increasingly seen as an important norm to maintain journalism's authority as information provider and to uphold the public's trust (Karlsson, 2011; Phillips, 2010b; Plaisance, 2007). Enabling the public to inspect the working process of journalists not only indicates the journalists' accountability, it also promotes the legitimacy of the news

production toward the public (Rennen, 2000). Attributing to sources—such as the news agency—is a prime example of the journalistic code of transparency. Yet dependency on agency copy is rarely publicly acknowledged by the agencies' clients (Boyer, 2011; Paterson, 2007; Quandt, 2008): Content is often subtly rewritten and either not attributed or erroneously attributed to a journalist, as demonstrated in various countries, including Greece (Doudaki & Spyridou, 2013; Saridou, Spyridou, & Veglis, 2017), Australia (Johnston, 2009), and the Netherlands (Scholten & Ruigrok, 2009). This study extends previous studies by comparing the transparency of print and online news. It is our expectation that:

H3: There is more attribution to the agency in print news than in online news.

Case Description: The Dutch News Landscape

The hypotheses are tested on a corpus of Dutch news articles. The Netherlands can be regarded as a prime example of a media system with a free press in which the print press sector follows free-market principles (although state subsidies do exist). Similar to other Northern and Central European nations, the Netherlands traditionally has relatively high levels of newspaper readership (Bakker & Scholten, 2011), even though it suffers from declining readership and declining revenues from advertising (Rutten & Slot, 2011). Consistent with findings in other national contexts (Lewis, Williams, Franklin, Thomas, et al., 2008; Paterson, 2007), the severe economic cutbacks in the Dutch print sector as observed over the past two decades have been linked to an increased reliance on the news agency (Welbers et al., 2016). Since news agency Geassocieerde Pers Diensten ceased to exist in 2013, Algemeen Nederlands Persbureau (ANP) is the only remaining news agency in the Netherlands (except for Novum Nieuws, which is owned by ANP). This market concentration makes the country a particularly suitable case for examining national news agency reliance. At the same time, given the Netherlands' similarity to other Western—in particular Northern European—media landscapes, we believe that our findings allow for inferences beyond the Dutch case.

Method

The impact of the agency on the news can be assessed in multiple ways. This article introduces an innovative automated tool to investigate two related levels: first, the level of the media agenda (i.e., What percentage of the news outlet's articles is initiated by agency copy?) and, second, on the content level (i.e., How similar is the content of the news article to the agency copy?).

Data

To assess these questions, an extensive data set has been collected: It includes all releases of the Netherlands' largest news agency, ANP, in 2014 as well as all news articles that appeared in both the print and online editions of the national newspapers *de Volkskrant*, *De Telegraaf*, and *Metro*.² These

² For the online news of *de Volkskrant* and *De Telegraaf*, only freely available content has been collected. While the majority of the articles are available, part of their content is behind a paywall. Our findings on the online editions of *de Volkskrant* and *De Telegraaf* are thus limited to their freely available content.

newspapers are, respectively, the largest quality, popular, and free newspapers in the country. Agency and print data have been retrieved from the LexisNexis archives. The online data were collected in real time by running a script on a server that queried the RSS feeds of the websites once per hour and downloaded the full text of the article. Finally, we have included the news site Nu.nl, which is the largest native online news organization in the Dutch online news market with an average of 8 million unique visitors per month ("NU.nl is het grootste", n.d.; Trilling & Schoenbach, 2015). Given our interest in the agency's impact on current affairs, some sections have been filtered out. This filtering is based on a close examination of the news sections as labeled by the titles. Filtered sections include opinion pieces, columns, editorials, cartoons, travel, show business, lifestyle, weather, and reviews of books, films, and music. Since our focus is on (inter)national news, regional sections have also been excluded. The final data set consists of 119,452 agency articles (equaling 327 articles per day) and 247,161 news articles. Table 1 provides an overview of the data set.

Table 1. Total Number of Articles in the Data Set.

Agency	<i>n</i>	Print	<i>n</i>	Online	<i>n</i>
ANP	119,452	<i>de Volkskrant</i>	25,191	Volkskrant.nl	49,471
		<i>De Telegraaf</i>	38,625	Telegraaf.nl	72,378
		<i>Metro</i>	11,618	Metronieuws.nl	34,503
				Nu.nl	15,375
Total	119,452		75,434		171,727

From Table 1, it is clear that for all three newspapers, the number of online news articles is substantially greater than the number of print articles. The popular *De Telegraaf* produced by far the most news articles, both in print as well as online: a daily average of, respectively, 124 and 198 articles, based on six print publication days and seven online publication days per week. The online-only medium Nu.nl published the fewest articles among online editions: $n = 15,375$ articles—or 42 articles per day. In terms of article length based on the number of words in the article (not depicted in the table), online articles are substantially shorter ($M = 194$) than print articles ($M = 329$).

Analysis

We developed a software tool to automatically compare print and online news content with agency content. In a first step, the tool parses and preprocesses the text (stop word and punctuation removal) and creates a term frequency-inverse document frequency representation of each text (for an explanation of such preprocessing steps, see Boumans & Trilling, 2015). In a second step, the tool calculates two measures: the intermedia agenda-setting ratio and the churnalism index. The calculation of both is described below. The tool is developed with open-source software: the Python libraries NumPy (version 1.10.4; Van Der Walt, Colbert, & Varoquaux, 2011) and SciPy (version 0.16.1; Van der Walt et al., 2011) and scikit-learn (version 0.17.0; Pedregosa et al., 2011).

Intermedia Agenda-Setting Ratio

We operationalize the intermediate agenda-setting ratio as the percentage of news articles that is initiated by agency copy. This ratio provides insight into the extent to which the newspapers' agendas are based on the news agency's agenda. Determining whether an article is initiated by an agency release is based on a measure called cosine distance, which indicates how similar two documents are likely to be in terms of their subject matter (Tan, Steinbach, & Kumar, 2006). Put more simply, it indicates the extent to which they share the same words. A cosine score can take any value between 0 and 1, depending on the degree of similar terms in the two texts. A score of 0 implies that two documents do not share any terms, and a score of 1 implies that the words in the two documents are identical (disregarding the word order).

A systematic manual analysis of a subset of the data showed that a similarity score of .65 and above indicates a content overlap that is too high to be reasonably explained by chance: The majority of the content is identical. This value therefore serves as threshold to determine whether or not there is a link between a media article and agency copy. When the cosine score is .65 or higher, the software reports the existence of a link. This threshold is similar to the threshold of .70 as proposed and validated by Welbers and colleagues (2016). The minor difference is likely due to the slightly different approach: while Welbers and colleagues only look at the nouns and proper names of texts, the current approach compares entire texts (stop words excluded). As an additional condition, we specified that the agency copy must be published on the same day, the day before, or two days before the news article. The formula of the ratio is as follows:

$$\text{Intermedia agenda-setting ratio} = n_{ia} / n_{ta}, \quad (1)$$

where n_{ia} = the number of media articles initiated by an agency release, and n_{ta} = the total number of media articles on that day. The following hypothetical example illustrates the agenda-setting ratio. Let us assume that on a certain day, the newspaper *De Telegraaf* publishes 40 articles ($n_{ta} = 40$). The software tool finds that 10 of those newspaper articles are based on agency releases that have been published in the three days before the newspaper article ($n_{ia} = 10$). The agenda-setting ratio is thus $10/40 = .25$, indicating that 25% of the *De Telegraaf's* coverage on that day is initiated by (based on) the agency releases.

Churnalism Index

The next step determines the extent to which the news article literally replicates agency content. At the core of the developed churnalism index is the Levenshtein distance (lev), a well-established measure in computer science and information theory that is often used in plagiarism detection tools. Lev is the minimum number of edits required to change one string into another. An edit can be the insertion, deletion, or substitution of a character. The more two texts differ from each other, the more edits are required to make the two texts identical, and thus the higher the lev score. When two texts are identical, the lev score is zero. For the purpose of this study, we are interested in the degree to which a news article is made up of agency copy. When a news article is a shortened replication of agency copy, the lev score equals the number of excess words of the agency text. Yet the act of merely cutting does not constitute a

journalistic contribution since there is no information added: The final product still entirely consists of agency copy and can be regarded as a prime example of the copy-and-paste practice that is at the core of the churnalism notion. Therefore, the measure controls for length difference. Furthermore, we want to analyze the *relative* effort a journalist has put into a media text. Twenty alterations of a source text that consists of only 30 words are not the same as 20 alterations of a 500-word text. We therefore also control for the length of the journalistic text. Formally, then, the formula for the churnalism index is the following:

$$\text{Churnalism index} = 1 - [\text{lev}_{a,b} - (\text{LI}_a - \text{LI}_b)] / \text{LI}_b, \quad (2)$$

where $\text{lev}_{a,b}$ = the Levenshtein distance between source text a and media text b , LI_a = the length indicator of the source text, and LI_b = the length indicator of the media text. By definition, the value of the churnalism index ranges from 0 to 1. The measure is inversed to facilitate interpretation: The higher the score, the greater the overlap (and thus the higher the degree of churnalism). A systematic manual analysis of a subset of the data reveals that when the churnalism index $> .7$, the two texts are nearly identical; a score of 0 indicates that, while there is some resemblance in terms of topic and word usage, the media text differs substantially from the source text.

Results

Our overall expectation is that online news media rely on agency content more than print news media do. The first dimension of this reliance considers the extent to which the news agenda is initiated by the agency (H1), and the second dimension considers the extent to which the content is verbatim agency copy (H2). Finally, we assess whether online titles attribute the agency less than the print titles do, as hypothesized (H3). We start by discussing some key findings that offer a first insight into the extent to which the national news agenda and content are based on agency copy.

In total, 119,555 news articles were published in 2014 that appear to be predominantly based on agency copy. This implies that, overall, nearly half of the news content (48.4%) is based on agency copy. The extent to which the agency exerts influence on the news can also be approached from the angle of the agency's output. Thus, we could investigate what proportion of the agency's total output is followed up on by the news media. It can be argued that high proportions are indicative of a greater dependency on the agency, while low proportions of followed-up agency copy indicate that media deliberately cherry-pick from the supply of content. As depicted in Table 1, the total number of agency articles in 2014 was 119,452. Results indicate that more than two of every five articles published by the agency are followed up on at least once (48,884 reproduced unique ANP articles/119,452 total ANP articles = 41%). On average, a followed-up article provides input for 1.76 news articles ($SD = 1.08$). When distinguishing between online and off-line editions, results indicate that the online media rely considerably more on the agency's agenda: 58% of the agency's content is picked up by at least one online news site, whereas only 17% of the agency's content is used by the print media. There thus appears to be a high demand for agency copy in the online news market: Over half of the agency's proposed items on the agenda are taken up by at least one online news provider. The next section examines the intermedia agenda-setting ratio to determine the extent to which this reproduced copy shapes the print and online news agendas.

Intermedia Agenda-Setting Ratio

For all three newspapers, the online edition is initiated by agency content more so than the print counterpart (see Figure 1). The ratio is highest for the online-only title Nu.nl. The ratio indicates that 75% of the news articles that appear on the website are based on agency copy. The news agenda of Nu.nl is thus predominantly driven by agency input. To compare, the ratio of the print titles is 16% for *de Volkskrant*, 12% for *De Telegraaf*, and 48% for *Metro*. An independent-samples *t* test was run to determine whether there are differences in the mean agenda-setting ratios of print and online news. Levene's test for equality of variances indicated heterogeneity of variances ($p = .00$), so a Welch *t* test was employed. Analysis confirms that online news is statistically significantly more often initiated by agency content than print news is, $M_{\text{online}} = .68$, $SD = .13$; $M_{\text{print}} = .23$, $SD = .18$, $M_{\text{diff}} = .44$, $t(1,537) = 64.82$, $p = .00$. Our first hypothesis is thus confirmed. Relating these outcomes to the findings that the online news providers reproduce 58% of the agency's total output, it becomes clear that the agency's agenda to a large degree *is* the online news agenda.

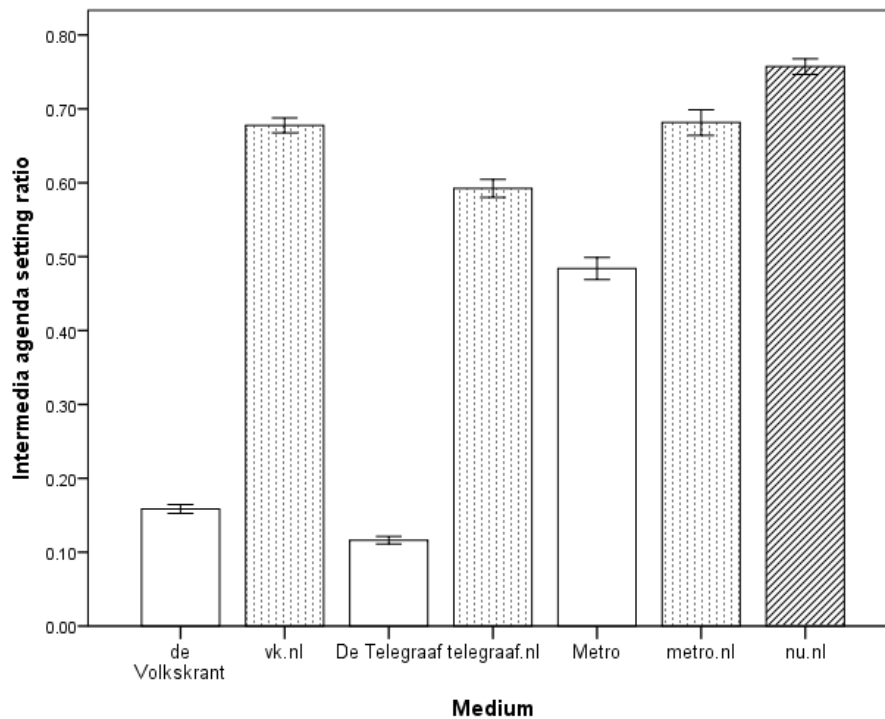


Figure 1. Intermedia agenda-setting ratio per title. Error bars represent standard errors.

Thus far the focus has been on the extent to which the agency shapes the news agenda. The second part of the analysis sketches a detailed insight into the extent to which agency copy is verbatim replicated.

Churnalism Index

The second hypothesis proposed that the content of online news media is more similar than the content of print newspapers to news agency copy. Figure 2 presents the average churnalism index score for all the news articles that have demonstrable overlap with agency copy.

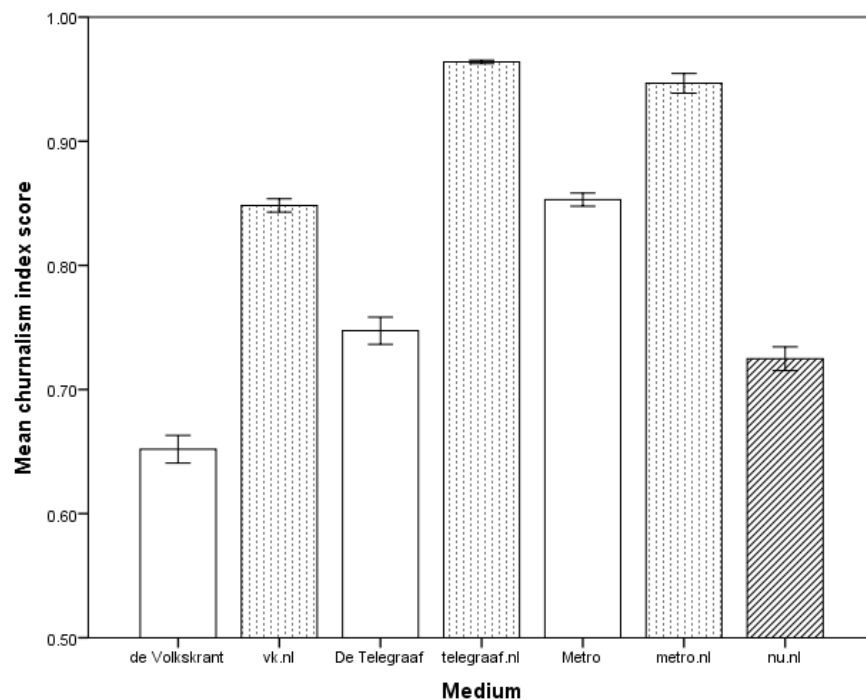


Figure 2. Mean churnalism index scores per medium. Error bars represent standard errors.

Articles in the online edition generally score higher than articles in the print edition on the churnalism index (see Figure 2). This is true for all three newspaper titles. In other words, online articles that are based on agency copy have been edited less than print news articles. To assess whether these differences are statistically significant, we performed an independent-sample *t* test. There was homogeneity of variances, as assessed by Levene's test for equality of variances ($p = .801$). Results confirm that the churnalism index score is higher for online news ($M = .87$, $SD = .11$) than for print news ($M = .74$, $SD = .12$), a statistically significant difference, $M_{diff} = .13$, $SE = .00$, $t(2,667) = 29.82$, $p = .00$.

The expectation that the content of online news media is more similar than the content of print newspapers to news agency copy (H2) is thus confirmed. It is worth noting the performance of online-only news provider Nu.nl: With a mean churnalism index score of .72 ($SD = .09$), it is slightly lower than the average print score; in fact, it is the lowest score after the print quality newspaper *de Volkskrant* ($M = .65$, $SD = .10$). Thus, while the vast majority of its news originates from the agency (Nu.nl has a ratio of .75), it rewrites and edits this input relatively more than the other media. In other words, it appears that the journalistic routine at Nu.nl can be described as repurposing agency copy. A final result worth highlighting is the difference within print newspapers: The pattern that emerges is in line with the reasoning that quality newspapers, due to larger journalistic capacities, are least inclined to verbatim reproduce agency copy, followed by the popular and free newspapers.

To illustrate what the various scores practically mean in terms of content overlap, we will briefly discuss the mean scores. A churnalism index score of .87 (mean score of the online titles) indicates that the agency copy is nearly verbatim taken up by the news outlet; typically, the agency text is shortened and at most one or two words are rephrased. An index score of .74 (mean score of the print titles) indicates that the agency text has been edited more, yet still rules out that the texts differ substantially from each other; the news article still consists of mostly agency content. Figure 3 demonstrates this with a comparison of an agency text and a newspaper article. The newspaper article (depicted on the right) has copied the majority of the agency's text, yet is slightly edited, as indicated by the headline and the first sentence. The article contains no other information than that provided by the agency.

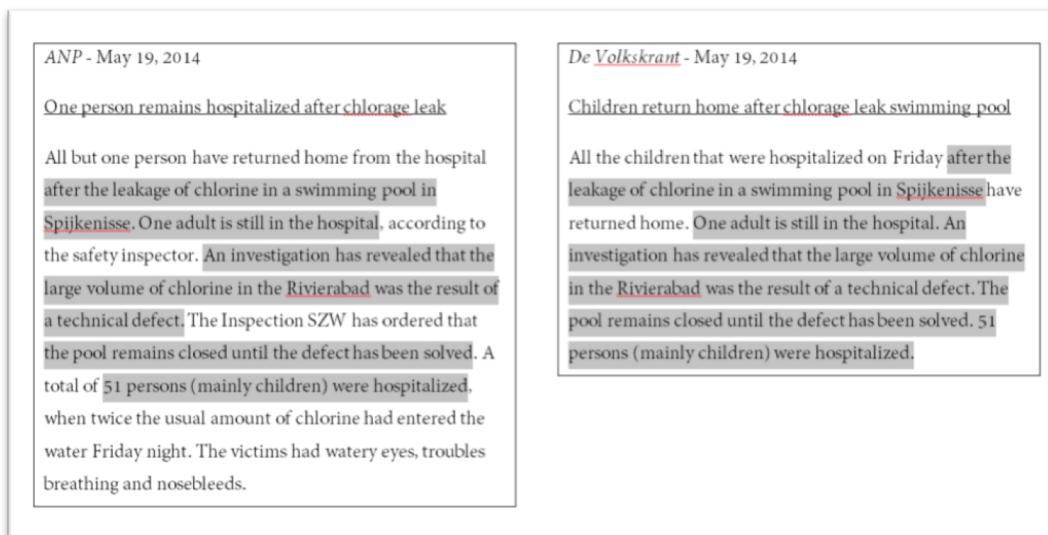


Figure 3. Example of two texts with a churnalism index score of 0.74. Identical phrases are highlighted.

Attribution

Our final research question concerns the extent to which titles attribute the agency as a source. Figure 4 provides an overview of the agency attribution ratio per title. The popular *De Telegraaf* attributes the agency neither in print nor online. Thus, readers of *De Telegraaf* are not informed that the information comes from an agency rather than from a newspaper journalist. It was our expectation that online titles would attribute less often than print titles (H3). For *de Volkskrant*, the data suggest a reverse pattern: Whereas the print version attributes the agency in 70%, the online version acknowledges the agency in 95% of the agency-based content. The print and online editions of the free daily *Metro* score comparably high (respectively, .94 and .93). Of all titles, Nu.nl scores highest on the attribution ratio: .96. Hypothesis 3, stating that print media attribute the agency more than online media, must thus be rejected. In fact, results from an independent-samples *t* test show that the opposite is true: Overall, the online titles perform better than the print titles on the attribution ratio, $M_{diff} = .19$, $SD = .02$, $t(2,184) = 10.32$, $p = .00$ (equal variance assumed), Levene's test statistic $p = .20$.

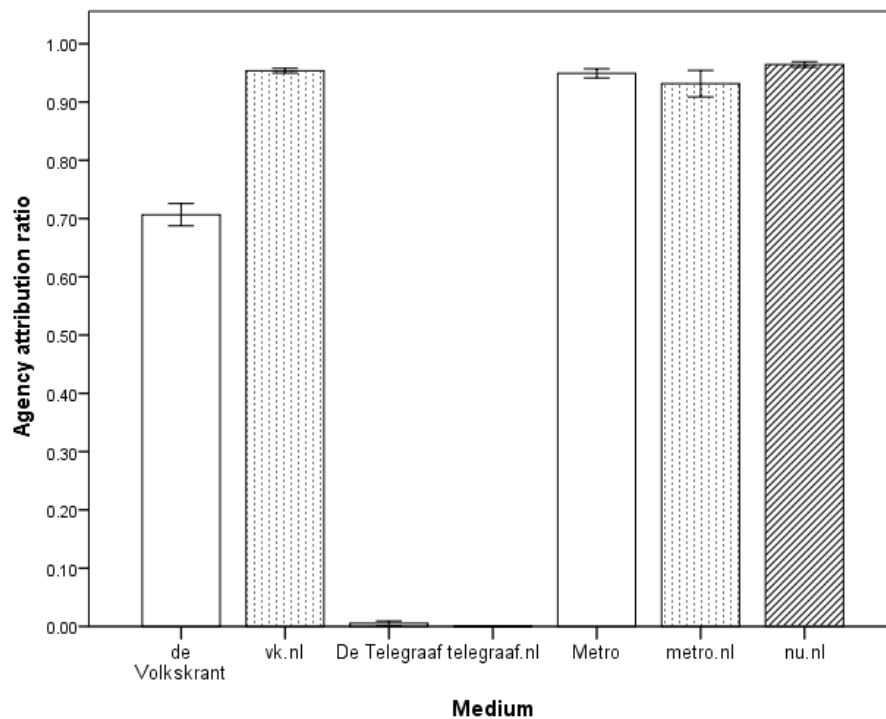


Figure 4. Agency attribution ratio. Online media are patterned, and error bars represent standard errors.

Conclusion

Informed by a unique data set, this study provides insight into the impact of news agencies on the shaping of the daily news. Analyses demonstrate that the agency provides input for almost a quarter of print news. However, this percentage is strongly influenced by the free outlet's high reliance on agency copy. For the quality and popular newspapers, the ratio is respectively a factor two and three lower than the ratio of the free outlet. The overwhelming majority of news in the paid newspapers does not contain (substantial amounts of) agency copy. The churnalism scores of print newspapers follow the expected pattern: lowest score for the quality paper, highest score for the free outlet. The reliance of online news on agency copy, however, presents a very different picture than for the print sector: 68% of the online news comes from the desk of the agency. For the primary online news provider, Nu.nl, this reliance on the agency is even higher: three out of every four articles are derived from agency copy. Combined with the finding that almost three out of every five agency articles end up as online news, it can be argued that the agency's agenda de facto is the online news agenda. Furthermore, the online content shows a great deal of overlap with the agency copy, as indicated by the churnalism index: The mean scores indicate that, in general, no value in the sense of new information (e.g., from an alternative source) is added.

While the distinction between online and print news can be somewhat diffuse—some news organizations have one newsroom to produce both print and online content—the findings presented here suggest that the two domains generally rely on different routines to create their products. The analysis of news articles of an entire year demonstrates that earlier expressed fears that “copy & paste becomes the basic principle” (Quandt, 2008, p. 729) for online news were not exaggerated. Although some minor editing takes place, the high churnalism index indicates that original reporting is scarce in online reporting. It is worth noting that online news providers do not hide that they are “on the agency's drip”: Apart from the popular one, all online outlets attribute consistently. The entire absence of attribution by the country's largest popular outlet, *De Telegraaf*, on the other hand, is disconcerting for those who value what has been labeled “a new norm within journalism” (Karlsson, 2011, p. 279), as transparency has been regarded. Indeed, transparency can be achieved in various ways; yet being open about agency reliance would be a relative straightforward one.

We emphasize that news agencies are high-quality sources and perform an important public service function, and there is nothing inherently wrong with the use of agency copy. Yet we believe an excessive reliance on the agency should be viewed with concern for at least two reasons. The first is that the omnipresence of agency content is at odds with the ideal of a pluralist media landscape. Research on international news coverage has consistently found that international news coverage is shaped by only two global news agencies, which results in a narrow account of global affairs (Boyd-Barrett & Rantanen, 1998; Paterson, 2001, 2011). This study shows, in a similar vein, that national news agencies dominate the domestic online news landscape, corroborating observations made in other national contexts (Paterson & Domingo, 2008; Quandt, 2008). What emerges is a picture of agencies not only providing the “raw material” but actually selecting and telling a considerable part of the daily news stories. This fuels concerns on the level of homogeneity and potential lack of diverse viewpoints in news—concerns that have been voiced by other scholars (Doyle, 2015; Fenton, 2010; Paterson, 2007, 2011). Although we know that international agencies tend to focus on a limited menu of news stories and news frames, favoring a small

group of elite nations, the empirical question of the actual or implied determinants of news coverage and news selection of national news agencies is yet to be decisively answered (Boyd-Barrett, 2010, p. 41). Since journalism's increasing dependency on news agency copy has been reported for several countries, the urgency of such research transcends the Dutch context.

The second reason for concern about the dominant role of the news agency follows from the agency's viability and the risk of a failing market of a commodity that is in essence a public good. Challenges such as falling revenues, pressure from news concerns, and problems with online piracy of the agency's content are reported for several national agencies (see Boyd-Barrett, 2010), but the Dutch agency is particularly vulnerable. This is partly because the ANP belongs to the small group (of around 15%) of national agencies worldwide that is not state supported (Boyd-Barrett, 2010), but is instead, for the most part, dependent on media clients. As documented for Scandinavian news agencies, the economic challenges are greater for agencies that receive no subsidies from the state (Rantanen & Boyd-Barrett, 2004). For the ANP, the increased pressures have resulted in multiple rounds of layoffs over the past decade (Rutten & Slot, 2011). As suggested in the case of the UK Press Association, an understaffed agency can lead to a news environment in which corporate and governmental public relations efforts are facilitated rather than challenged (Davies, 2008; Lewis, Williams & Franklin, 2008). Although, like many other agencies, the ANP has adopted a new business model by jumping on the convergence bandwagon and reinventing itself as "partner in the entire communication cycle" ("ANP-communicatiecyclus," n.d.) to the business sector, the agency is still not profitable (Brandenburg-Van de Ven, 2015). In 2011, the ANP's president warned that increasing market pressures pose a threat to the quality of the agency's output (Rutten & Slot, 2011, p. 25). The exceptionally concentrated Dutch newspaper market is an additional problem for the ANP, because it places the agency in a weak negotiation position against its main clients, the news organizations. The agency has been cooperatively owned by these organizations since its creation in 1934, but it was privatized in 2003. This development structurally altered the relation between the agency and its clients: Solidarity has been replaced by pragmatic and commercial interests (Rutten & Slot, 2011). As in several other countries, including the United States, Canada, France, Germany, and the United Kingdom (Boyd-Barrett, 2010; Paterson, 2010), the largest news organizations consider in-sourcing the wire service. Although historically most such efforts have not ended successfully for news organizations (Boyd-Barrett, 2010), given the small size of the Dutch news market, such an initiative could nonetheless mean the end of the ANP (Rutten & Slot, 2011).

While the market conditions in which the Dutch agency operates are particularly tough and some agencies have managed to—at least commercially—successfully adapt to the technological challenges and changed circumstances on the market (Nogué Regàs, 2010; Paterson, 2010), there are many more national agencies that face a similar crisis of survival and identity (Boyd-Barrett, 2010; Ellis, 2007; Rantanen & Boyd-Barrett, 2004). In the specific case of the Dutch national agency, it is evident from this study that the consequences of a collapsing agency are immense for particularly the online news landscape: It would, to a large extent, be depleted of content. Furthermore, a bankruptcy of the country's only national agency would mean the disappearance of the market where Dutch news organizations can buy (national) news services. The largest news organizations will be able to in-source this service, but for smaller news organizations it would be detrimental.

The measurement techniques presented here do warrant some additional discussion. One point to consider is the threshold that was chosen to determine whether a link exists between an agency text and a news article. The threshold is carefully decided upon after manual analysis of hundreds of links. To avoid including spurious links—which could occur when an agency text and a newspaper article both use the same paragraph or quote from a press release or statement—it is deliberately set at a moderately high value. While this gives the assurance that all detected links are indeed indicative of newspapers' usage of agency copy rather than coincidence, a consequence is that articles that to a lesser degree rely on agency copy are not included in the agenda-setting ratio. The ratio thus provides insight into the proportion of news articles that are *largely* based on agency copy rather than the full extent to which newspapers rely on agency copy. In other words, the numbers presented here are conservative estimates; the actual impact of the agency is even higher if we take weaker links into account. With this in mind, the already high ratios for the online titles are even more striking.

By only considering agency copy that appeared *before* a newspaper article, we reduced a possible alternative explanation for our findings: that the agency and the newspaper independently of each other replicate the same source text, such as a press release. Although this possibility cannot be ruled out entirely, there are two arguments against this alternative explanation. First, even though public relations sources are increasingly important for journalists—some speak of a “Niagara of PR-propaganda” (Moloney, 2006, p. 2) falling on journalists (see also Davis, 2000; Jackson & Moloney, 2016)—various studies have shown that when a press release is relied on as input for a news article, journalists generally transform and augment the content (Maat & de Jong, 2013; Reich, 2010; Tenenboim-Weinblatt, & Baden, 2016). Second, another study that is based on the same approach and same news media as this study demonstrated that literal copy-and-paste practices of press releases is virtually nonexistent in the Dutch mainstream press: Of the 6,147 agency articles and 19,985 newspaper articles about organizations, one and three articles, respectively, were found to be replications of press releases of those organizations (Boumans, 2017). It thus appears highly improbable that the scenario of independent source replication forms a structural explanation for the overlap found in agency and newspaper content.

Overall, the measures have proved useful indicators and facilitate systematic comparative research on both agenda setting and churnalism. The ability to analyze large volumes of data allows for cross-national as well as longitudinal research. Furthermore, the proposed churnalism measure allows for a systematic assessment of the degree to which subsidized content from different suppliers makes it into the news.

In conclusion, our findings provide statistical support for earlier claims of news agencies' determining impact on the news and make a strong case for a systematic scrutiny of their role. The approach presented here can foster such research. While the unavailability of the print data in digital archives may pose a problem for research in certain countries, retrieving news (both online and print) in real time is quite possible. Not only researchers interested in agenda-setting dynamics should be alerted: The fact that the most important online news providers largely present news that comes from one and the same source is problematic for anyone concerned with news diversity. Research could advance in at least two directions: (a) toward gaining a better understanding of the news production process of the agencies, including the nature and quality of their output; and (b) toward greater insight into the journalistic

routines with respect to agency copy. To accomplish the latter, different approaches could assess how agencies influence news selection processes within their client media. Concerning the first, sourcing and verification practices are a particularly important area of interest. While the factors that determine agency selection are still poorly understood by researchers (Boyd-Barrett, 2010; Paterson, 2010), concerns have been expressed that the agencies' strategies to cope with economic pressures will be detrimental for the quality of their output (Davies, 2008; Wilke, 2010). The findings presented in this study suggest it is time for a rigorous assessment of these concerns.

References

- ANP Communicatiecyclus. [ANP Communication cycle]. (n.d.). Retrieved from <https://www.anp.nl/anp-communicatiecyclus>
- Bakker, P., & Scholten, O. (2011). *Communicatiekaart van Nederland: Overzicht van media en communicatie* [Communication map of the Netherlands: An overview of media and communication]. Alphen aan de Rijn, Netherlands: Wolters Kluwer.
- Boumans, J. (2017). Subsidizing the news? Organizational press releases' influence on news media's agenda and content. *Journalism Studies*, 1–19. doi:10.1080/1461670X.2017.1338154
- Boumans, J., & Trilling, D. (2015). Taking stock of the toolkit: An overview of relevant automated content analysis approaches and techniques for digital journalism scholars. *Digital Journalism*, 4(1), 8–23. doi:10.1080/21670811.2015.1096598
- Boyd-Barrett, O. (1980). *The international news agencies (Vol. 13)*. London, UK: SAGE Publications.
- Boyd-Barrett, O. (2000). National and international news agencies: Issues of crisis and realignment. *Gazette*, 62(1), 5–18. doi:10.1177/0016549200062001001
- Boyd-Barrett, O. (Ed.). (2010). *News agencies in the turbulent era of the Internet*. Barcelona, Spain: Generalitat de Catalunya/Col·leccio Lexikon.
- Boyd-Barrett, O., & Rantanen, T. (1998). *The globalization of news*. London, UK: SAGE Publications.
- Boyer, D. (2011). News agency and news mediation in the digital era. *Social Anthropology*, 19(1), 6–22. doi:10.1111/j.1469-8676.2010.00135.x
- Brandenburg-van de Ven, T. (2015, May 18). Verlies voor ANP. [Negative result for ANP]. Retrieved from <https://www.villamedia.nl/artikel/verlies-voor-ANP>
- Davies, N. (2008). *Flat earth news*. London, UK: Chatto & Windus.

- Davis, A. (2000). Public relations, news production and changing patterns of source access in the British national media. *Media, Culture & Society*, 22(1), 39–59. doi:10.1177/016344300022001003
- Doudaki, V., & Spyridou, L. (2013). Print and online news: Remediation practices in content and form. *Journalism Studies*, 14(6), 907–925. doi:10.1080/1461670x.2012.746860
- Doyle, G. (2015). Multi-platform media and the miracle of the loaves and fishes. *Journal of Media Business Studies*, 12(1), 49–65. doi:10.1080/16522354.2015.1027113
- Ellis, G. (2007). Rewired: How competition ended 125 years of cooperative national news coverage in New Zealand. *Political Science*, 59(2), 33–49. doi:10.1177/003231870705900204
- Esperidião, M. C. (2011). Invisible giants in broadcast journalism: News agencies and the global news ecosystem. *Brazilian Journalism Research*, 7(1), 104–127.
- Fenton, N. (2010). *New media, old news: Journalism & democracy in the digital age*. London, UK: SAGE Publications.
- Forde, S., & Johnston, J. (2013). The news triumvirate. *Journalism Studies*, 14(1), 113–129. doi:10.1080/1461670x.2012.679859
- Harcup, T. (2014). *A dictionary of journalism*. Oxford, UK: Oxford University Press.
- Hijmans, E., Schafraad, P., Buijs, K., & d’Haenens, L. (2011). Wie schrijft ons nieuws? [Who writes our news?]. *Tijdschrift voor Communicatiewetenschap*, 37(2), 77–91.
- Horvit, B. (2006). International news agencies and the war debate of 2003. *International Communication Gazette*, 68(5–6), 427–447. doi:10.1177/1748048506068722
- Humprecht, E., & Büchel, F. (2013). More of the same or marketplace of opinions? A cross-national comparison of diversity in online news reporting. *The International Journal of Press/Politics*, 18(4), 436–461. doi:10.1177/1940161213497595
- Jackson, D., & Moloney, K. (2016). Inside churnalism: PR, journalism and power relationships in flux. *Journalism Studies*, 17(6), 763–780. doi:10.1080/1461670x.2015.1017597
- Johnston, J. (2009). “Not wrong for long”: The role and penetration of news wire agencies in the 24/7 landscape. *Global Media Journal—Australian Edition*, 3(1), 1–16.
- Johnston, J., & Forde, S. (2011). The silent partner: News agencies and 21st century news. *International Journal of Communication*, 5, 195–214.

- Karlsson, M. (2011). The immediacy of online news, the visibility of journalistic processes and a restructuring of journalistic authority. *Journalism, 12*(3), 279–295.
doi:10.1177/1464884910388223
- Klinenberg, E. (2005). Convergence: News production in a digital age. *Annals of the American Academy of Political and Social Science, 597*(1), 48–64.
- Lewis, J., Williams, A., & Franklin, B. (2008). A compromised fourth estate? UK news journalism, public relations and news sources. *Journalism Studies, 9*(1), 1–20. doi:10.1080/14616700701767974
- Lewis, J., Williams, A., Franklin, B., Thomas, J., & Mosdell, N. (2008). The quality and independence of British journalism. Cardiff, UK: Cardiff University.
- Livingston, S., & Bennett, W.L. (2003). Gatekeeping, indexing, and live-event news: Is technology altering the construction of news? *Political Communication, 20*(4), 363–380.
doi:10.1080/10584600390244121
- Maat, H., & de Jong, C. (2013). How newspaper journalists reframe product press release information. *Journalism, 14*(3), 348–371. doi:10.1177/1464884912448914
- Manning, P. (2008). The Press Association and News Agency sources. In B. Franklin (Ed.), *Pulling newspapers apart* (pp. 247–256). London, UK: Routledge.
- McChesney, R. (2003). The problem of journalism: A political economic contribution to an explanation of the crisis in contemporary U.S. journalism. *Journalism Studies, 4*(3), 307–309.
doi:10.1080/14616700306492
- McChesney, R. (2008). *The political economy of media: Enduring issues, emerging dilemmas*. New York, NY: Monthly Review Press.
- Mitchelstein, E., & Boczkowski, P. J. (2010). Online news consumption research: An assessment of past work and an agenda for the future. *New Media & Society, 12*(7), 1085–1102.
doi:10.1177/1461444809350193
- Moloney, K. (2006). *Rethinking public relations: PR propaganda and democracy*. London, UK: Routledge.
- Nogué Regàs, A. (2010). The Agència Catalana de Notícies (ACN): Tenth anniversary of an innovative model. In O. Boyd-Barrett (Ed.), *National news agencies in the turbulent era of the Internet* (pp. 45–66). Barcelona, Spain: Generalitat de Catalunya/Col·leccio Lexikon.
- Nu.nl. Het grootste digitale nieuwsmerk van Nederland. [Nu.nl. The largest digital news brand of the Netherlands]. n.d. Retrieved from <https://www.sanoma.nl/brand/nu-nl/>

- Paterson, C. (2001). Media imperialism revisited: The global public sphere and the news agency agenda. In S. Hjarvard (Ed.), *News in a globalized society* (pp. 77–92). Gothenburg, Sweden: Nordicom.
- Paterson, C. (2007). International news on the Internet: Why more is less. *Ethical Space: International Journal of Communication Ethics*, 4(1/2), 57–66.
- Paterson, C. (2010). Changing times: The move online and the UK's Press Association. In O. Boyd-Barrett (Ed.), *National news agencies in the turbulent era of the Internet* (pp. 225–244). Barcelona, Spain: Generalitat de Catalunya/Col·leccio Lexikon.
- Paterson, C. (2011). *The international television news agencies: The world from London*. New York, NY: Peter Lang.
- Paterson, C., & Domingo, D. (Eds.). (2008). *Making online news: The ethnography of new media production*. New York, NY: Peter Lang.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Vanderplas, J. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12(Oct), 2825–2830.
- Phillips, A. (2010a). Old sources: New bottles. In N. Fenton (Ed.), *New media, old news. Journalism and democracy in the digital age* (pp. 87–102). London, UK: SAGE Publications.
- Phillips, A. (2010b). Transparency and the new ethics of journalism. *Journalism Practice*, 4(3), 373–382. doi:10.1080/17512781003642972
- Plaisance, P. (2007). Transparency: An assessment of the Kantian roots of a key element in media ethics practice. *Journal of Mass Media Ethics*, 22(2–3), 187–207. doi:10.1080/08900520701315855
- Powers, M., & Benson, R. (2014). Is the Internet homogenizing or diversifying the news? External pluralism in the U.S., Danish, and French press. *International Journal of Press/Politics*, 19(2), 246–265. doi:10.1177/1940161213519680
- Quandt, T. (2008). (No) news on the World Wide Web? A comparative content analysis of news in Europe and the United States. *Journalism Studies*, 9(5), 717–738. doi:10.1080/14616700802207664
- Rantanen, T., & Boyd-Barrett, O. (2004). Global and national news agencies: The unstable nexus. In A. S. de Beer, S. Arnold, & J. C. Merrill (Eds.), *Global journalism: Topical issues and media systems* (pp. 35–49). Boston, MA: Allyn & Bacon.
- Reich, Z. (2010). Measuring the impact of PR on published news in increasingly fragmented news environments: A multifaceted approach. *Journalism Studies*, 11(6), 799–816. doi:10.1080/14616701003760550

- Rennen, A. (2000). *Journalistiek als kwestie van bronnen: Ontwikkeling en toepassing van een bron-georiënteerde benadering van journalistiek* [Journalism as a matter of sources: Development and application of a source-oriented approach of journalism]. Delft, Netherlands: Eburon.
- Rutten, P., & Slot, M. (2011). Zijn de persbureaus te verslaan? De positie van Nederlandse persbureaus in de nieuwsketen [Can the press agencies be beaten? The position of Dutch press agencies in the news chain]. Retrieved from http://www.tno.nl/content.cfm?context=thema&content=inno_publicatie&laag1=897&laag2=919&item_id=860
- Saridou, T., Spyridou, L., & Veglis, A. (2017). Churnalism on the rise? Assessing convergence effects on editorial practices. *Digital Journalism*, 5(8), 1006–1024. doi:10.1080/21670811.2017.1342209
- Scholten, O., & Ruigrok, N. (2009). Bronnen in het nieuws. Een onderzoek naar ANP berichten in nieuws en achtergrondinformatie in Nederlandse dagbladen 2006–2008 [Sources in the news. A study of ANP-content in news and background information in Dutch newspapers 2006–2008]. In W. Jongbloed, E. Lauf, & R. Negenborn (Eds.), *Mediamonitor. Analyse en verdieping #1* [Media monitor. Analysis and deepening] (pp. 37–62). Hilversum, Netherlands: Commissariaat voor de Media.
- Sjøvaag, H. (2014). Homogenisation or Differentiation?, *Journalism Studies*, 15(5), 511–521. doi:10.1080/1461670X.2014.885275
- Tan, P., Steinbach, M., & Kumar, V. (2006). *Introduction to data mining*. Boston, MA: Pearson Education.
- Tenenboim-Weinblatt, K., & Baden, C. (2016). Journalistic transformation: How source texts are turned into news stories. *Journalism*, 1–19. doi:10.1177/1464884916667873
- Tijdelijke Commissie Innovatie en Toekomst Pers. (2009). *Adviesrapport De volgende editie* [Advisory report "The next edition"]. Den Haag, Netherlands: Ministerie van Onderwijs, Cultuur En Wetenschap.
- Trilling, D., & Schoenbach, K. (2015). Investigating people's news diets: How online news users use offline news. *Communications: The European Journal of Communication Research*, 40(1), 67–91. doi:10.1515/commun-2014-0028
- Van Leuven, S., Deprez, A., & Raeymaeckers, K. (2014). Towards more balanced news access? A study on the impact of cost-cutting and Web 2.0 on the mediated public sphere. *Journalism*, 15(7), 850–867. doi:10.1177/1464884913501837
- Vliegthart, R., & Walgrave, S. (2008). The contingency of intermedia agenda setting: A longitudinal study in Belgium. *Journalism and Mass Communication Quarterly*, 85(4), 860–877. doi:10.1177/107769900808500409

- Walt, S. van der, Colbert, S. C., & Varoquaux, G. (2011). The NumPy array: a structure for efficient numerical computation. *Computing in Science & Engineering*, 13(2), 22–30.
doi:10.1109/MCSE.2011.37
- Weaver, D., & Bimber, B. (2008). Finding news stories: A comparison of searches using LexisNexis and Google News. *Journalism and Mass Communication Quarterly*, 85(3), 515–530.
doi:10.1177/107769900808500303
- Welbers, K., Van Atteveldt, W., Kleinnijenhuis, J., & Ruigrok, N. (2016). A gatekeeper among gatekeepers: News agency influence in print and online newspapers in the Netherlands. *Journalism Studies*, 1–19. doi:10.1080/1461670x.2016.1190663
- Wilke, J. (2010). DPA and the news agencies of Germany. In O. Boyd-Barrett (Ed.), *National news agencies in the turbulent era of the Internet* (pp. 139–156). Barcelona, Spain: Generalitat de Catalunya/Col·leccio Lexikon.