City Research Online

City St George's, University of London

Citation: Bastos, M. T. & Zago, G. (2013). Tweeting News Articles Readership and News Sections in Europe and the Americas. SAGE Open, 3(3), pp. 1-18. doi: 10.1177/2158244013502496

This is the published version of the paper.

This version of the publication may differ from the final published version. To cite this item please consult the publisher's version.

Permanent repository link: https://openaccess.city.ac.uk/id/eprint/16383/

Link to published version: https://doi.org/10.1177/2158244013502496

Copyright and Reuse: Copyright and Moral Rights remain with the author(s) and/or copyright holders. Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge, unless otherwise indicated, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way. For full details of reuse please refer to City Research Online policy.

City Research Online: http://openaccess.city.ac.uk/ publications@citystgeorges.ac.uk

SAGE Open
http://sgo.sagepub.com/

Tweeting News Articles: Readership and News Sections in Europe and the Americas

Marco Toledo Bastos and Gabriela Zago

SAGE Open 2013 3:
DOI: 10.1177/2158244013502496

The online version of this article can be found at:
http://sgo.sagepub.com/content/3/3/2158244013502496

Published by:
SAGE
http://www.sagepublications.com

SAGE open

Additional services and information for *SAGE Open* can be found at:

Email Alerts: http://sgo.sagepub.com/cgi/alerts

Subscriptions: http://sgo.sagepub.com/subscriptions

Reprints: http://www.sagepub.com/journalsReprints.nav

© 2013 the Author(s). This article has been published under the terms of the Creative Commons Attribution License. Without requesting permission from the Author or SAGE, you may further copy, distribute, transmit, and adapt the article, with the condition that the Author and SAGE Open are in each case credited as the source of the article.

Downloaded from sgo.sagepub.com by Marco Bastos on September 25, 2013

SAGE Open
July-September 2013: 1–18
© The Author(s) 2013
DOI: 10.1177/2158244013502496
sgo.sagepub.com
SAGE

Tweeting News Articles: Readership and News Sections in Europe and the Americas

Marco Toledo Bastos[1] and Gabriela Zago[2]

Abstract

In this article we investigate the impact of social media readership to the editorial profile of newspapers. We analyze tweets containing links to news articles from eight of the largest national newspapers in the United States, United Kingdom, Spain, Brazil, and Germany. The data collection follows the first two weeks of October 2012 and includes 2,842,699 tweets with links to news articles. Twitter-shortened links were resolved using a three-pass routine and assigned to 1 of the 21 newspaper sections. We found the concentration of links to news articles posted by top users to be lower than reported in the literature and the strategy of relaying headlines on Twitter via automatic news aggregators (feeds) to be inefficient. The results of this investigation show which sections of a newspaper are the most and least read by readers in different parts of the world, with German readers placing greater emphasis on Politics and Economy; Brazilians on Sports and Arts; Spaniards on Local and National news; Britons and Americans on Opinion and World news. We also found that German and Spanish readers are more likely to read multiple national newspapers, while British readers more often resort to foreign sources of news. The results confirm that feedback to news items from a large user base is pivotal for the replication of content and that newspapers and news items can be clustered according to the editorial profile and principles of newsworthiness inherited from legacy media. The results of this investigation shed light onto the networked architecture of journalism that increasingly depends on readership agency.

Keywords

Twitter, newspapers, journalism, news sections, news items, readership

Introduction

During the first decade of the 21st century social media became a driving force for reporting and distributing news articles. Readers populated social media with recommendations and suggestions for news articles, offering a view into audience agency and giving newsmakers an opportunity to understand how readers react to news articles. Social networking sites allowed readers to engage with each other and share stories. News sites introduced widgets, small applications that can be installed into websites to link back to news outlet content. Social bookmarking tools empowered users to rate news content and amplified their preferences for individual stories.

During this period, news outlets turned to social networking sites to compete in a fast-changing media landscape. Traditional newspapers developed strategies and experimented with paywall models to monetize the circulation of news articles. Following the *Wall Street Journal, The New York Times* and German publisher Axel Springer AG introduced a metered paywall for their websites of daily national newspapers. British News International protected the content of The Times and The Sunday Times behind an "iron-curtain," and a number of Brazilian newspapers implemented metered paywalls to access the newspapers' content.

These attempts to monetize news followed market-based strategies to capture attention on social networks. A number of news outlets appointed social media editors and integrated print and digital operations. The evolving landscape of news delivery compelled organizations to invest in Research and Development and to investigate innovative business solutions beyond the industry domain. In this period, a number of news outlets—notably *The Guardian, The New York Times, USA Today*, and National Public Radio (NPR)—deployed publicly available Application Programming Interfaces (API) that allow users to programmatically access data using configurable parameters (Aitamurto & Lewis, 2013).

Twitter increasingly became a channel for breaking news stories (Ovide, 2009) and was associated with the diffusion

[1]University of São Paulo, Brazil
[2]Federal University of Rio Grande do Sul, Brazil

Corresponding Author:
Marco Toledo Bastos, University of São Paulo, Av. Prof. Lúcio Martins Rodrigues, 443 - Bloco A, sala 41. 05508-900, São Paulo, Brazil.
Email: marco@toledobastos.com

of hard news topics (Conover et al., 2011; Larsson & Moe, 2012) and with the exponential growth in infotainment news (Hargittai & Litt, 2011). As social media became a central channel for the distribution of news articles across the Internet, Twitter co-founder Biz Stone announced the company's ambition to create a "Twitter News Service" for major news organizations (Bastos, Travitzki, & Puschmann, 2012; Cowan, 2010). The announcement was in line with a series of modifications to emphasize Twitter's news and information-network strategy, and followed early attempts to build a news processing system that could harness the platform's potential for real-time news propagation (Sankaranarayanan, Samet, Teitler, Lieberman, & Sperling, 2009).

In this paper, we investigate how circulation and classification of news are impacted by the growing use of social networks to distribute news articles. Although online news has grown rapidly to become a major source of news in Western countries (Pew Research Center, 2012; Reuters Institute for the Study of Journalism [RISJ], 2013), research focusing on online news is largely informed by the rationale set by print media and seldom evaluates the impact of social media to the building blocks of newsmaking. We expect the results of this investigation to shed light onto the production of news that evolved into a live scheme of news reporting. We will show how the relative distribution of news articles per section changes across countries and why the difference between broadsheet and tabloid newspapers created by legacy media organizations is still appropriate.

Previous Work

The key factors governing the newsworthiness of information have changed considerably since Otto Groth (1928) described the newspaper qualifications and article attributes and Galtung and Ruge (1965) identified the 13 factors governing newsworthiness. Tunstall (1971) commented that because Galtung and Ruge's dataset was restricted to the coverage of international crises, the authors ignored day-to-day coverage of lesser, domestic and mundane news. Seizing on Galtung and Ruge's seminal work, journalism research have periodically revisited the principles of newsworthiness. Shoemaker and Reese (1996) offered a historical overview of the theory of newsworthiness and established that the process of selecting which events are covered by media depended not only on the topic, but also on the presentation of the topic and the available information.

The reevaluation of Galtung and Ruge's research on the factors' driving newsworthiness led to a general consensus that the context of print media is one of increasing editorial emphasis on entertainment (Franklin, 1997). Harcup and O'Neill (2001) commented on Franklin's work and pointed out that no contemporary set of news values can be complete without the entertainment factor. The authors offered a revised version of Galtung and Ruge's original set of factors into what Harcup and O'Neill called Newspaper Agenda. The criteria that govern newsworthiness became an important item on the agenda of journalism research and a number of empirical studies examined the predictability of news coverage in western industrialized (Kepplinger & Ehmig, 2006; Pritchard & Hughes, 1997) and developing countries (Schwarz, 2006).

The organization of news items according to sections is a legacy from print media that is subject to change due to audience interference on newsroom decisions. Print newspapers have traditionally operated with preprint templates dedicated to the news sections of the paper. These preprinted sections were often mass-produced and stacked in suitable storage areas prior to the offset printing of current news (Fernandez-Rana, Simon, & Peserik, 1974; Pérez-Peña, 2008). Schudson (2011) commented that the distribution of items per section have changed from mid-20th-century news reporting, focused on political and socioeconomic issues, to infotainment news covering celebrities' personal lives and showbiz events. But despite the substantial literature on online news as a source of information, and on news sections on print press, research focusing on the effects of audience agency on the distribution of items per section is still forthcoming.

The early literature on digital news was mostly focused on investigating whether information retrieval was more efficient in digital platforms, thereby allowing for a relative increase in user's information consumption and news retention. In the late nineties, a group of Scandinavian researchers visited three news organizations over a period of 3 years. The researchers investigated structural changes in the production of news due to digital platforms and reported differences that set news websites apart from print media (Eriksen & Ihlstrom, 2000). Still according to this early research, news websites focus on hard news and foster the continuous streaming of news similar to live reporting. During the same period, Dimitrova, Connolly-Ahern, Williams, Kaid, and Reid (2003) looked at how online newspapers used hyperlinks to refer readers to outside websites and suggested that the gatekeeping role of online editors had remained strong.

There is also a sizable literature on the impact of online news to print newspapers. Filistrucchi, Fiesolana, and Domenico (2005) investigated the effect of websites of daily newspapers in Italy and reported a negative impact on sales of local newspapers. George (2008) studied the effects of the Internet to the audience of traditional American newspapers and found evidence that the Internet attracts younger and educated individuals away from daily newspapers and the opposite effect (higher newspaper circulation) among African Americans and Hispanics. Gentzkow (2007) analyzed the Washington, D.C., market with a model that evaluates complement consumption instead of competition between existing products. Gentzkow found that print and online newspapers are strong complements, even though online newspaper reduced print readership by 27,000 per day.

Research focused on readership of news sections across different media platforms is relatively scarce. D'Haenens,

Jankowski, and Heuvelman (2004) investigated the variations in readership between digital and print versions of two Dutch newspapers and reported no significant differences in the consumption of news in print and online platforms. However, the researchers noted a significant difference in the amount of time spent reading each news sections in print and online newspapers and found that more time was spent reading the online versions of the newspapers across all news sections, except for sports and local news, and that online readers recall more international news and less sports than readers of the print version. The research also found a much higher volume of local news and a higher volume of news articles overall in the print edition of the newspapers.

Quandt (2008) compared the overall distribution of articles per news sections on online news websites with patterns from content analyses of press and TV coverage (Shoemaker & Cohen, 2012). The researcher found a similar pattern across the news ecosystem, with emphasis on national politics and economy, followed by human-interest stories, international politics, crimes, sport, and culture. The investigation reported that online journalism—perhaps contrary to expectation—was fairly conventional in regard to the distribution of news items per news section (or topic categories). Although the researchers found striking similarities between websites inside a country, they also reported remarkable differences from country to country.

Still according to Quandt (2008), German websites featured less national politics and more economy, sports, and culture; British websites featured an average number of items about politics, but far more human-interest stories and "social affairs" in comparison with websites in other countries; French websites featured a high number of world news, but mostly focused on politics, with 60% of Le Monde's articles and 50% of Figaro's news articles being classified as politics; U.S. websites presented a high volume of news articles on crime-related stories and a higher-than-average volume of political news articles, with about 50% of the content being classified as politics; Russian websites covered mostly politics, economy, and crime. The results indicated that news websites were not nearly as international as the researcher expected, with content being limited by the traditional, national, and linguistic interests of the readers.

Recent studies have addressed the newspapers' audience beyond the paradigm of production associated with unidirectional information. Aitamurto and Lewis (2013) investigated the impact of open APIs at four news organizations and reported the change in news consumption habits as audiences become more fragmented and accustomed to gathering news from several sources rather than a single destination site. Mitchelstein and Boczkowski (2009, 2010) reviewed the research on online news production and described a prevailing preference for theoretical and methodological proposals that evolved from the American literature on the topic, and also commented on the scarce availability of research that considered events taking place in other parts of the world. The researchers advocate the need for investigations that look at phenomena across different locales.

As newsrooms become more dependent on networking technology, and as journalists increasingly use social media to reach the public (RISJ, 2013), researchers have started to investigate the impact of social media to the news ecosystem. Hong (2012) examined the relationship between news organizations' adoption of social media and online readership and reported a positive correlation between newspapers' adoption of social media and increase in online readership. Messner, Linke, and Eford (2011) used content analysis to investigate the adoption of Twitter by American news outlets and reported that news organizations failed to use Twitter as a community-building tool. Common to these recent studies is the focus on social media agency, with readers engaging in instantaneous responses to articles or events and imposing challenges to the scholarly division between studies of production and consumption tied to the unidirectional information architecture.

Objectives

In this article, we investigate how major print newspapers perform in social media across five countries, taking into consideration audience agency and the integration of multiple media ecosystems. We investigate how the distribution of news articles changes across countries and whether traditional models of print media endure in social networks. We hypothesized that the editorial profile of print newspapers is reflected on social media and can be confirmed using cluster analysis of the relative volume of news articles covering Politics, Economy, Sports, Opinion, and National or World news. We also expect the clustering algorithms to identify newspapers dedicated to hard news stories and papers committed to soft news topics, particularly Entertainment, Crime, Sports, Lifestyle, Celebrities, and events triggered by the public's curiosity.

Recent studies have explored the correlation between the distribution of news articles per section streamed on Twitter and the distribution of articles per section published on quality papers (Bastos, in press). The results indicated that Twitter is the only social network to present a statistically significant correlation with the distribution of news items per sections on *The Guardian* and *The New York Times* ($r = .89$ and .68, respectively, $p < .001$). Based on these recent findings, we expect the diversity of links to news articles that appeared on Twitter stream to be a good predictor of the editorial profile of online newspapers.

Although we cannot determine the consistency of the criteria used to assign items to news sections across news outlets, this can be partially controlled for by analyzing only quality papers that rely on a more uniform baseline of news editors' decisions. For that reason, we have excluded popular tabloids from this research and focused on quality papers with editorial styles that encompass the extremes of the

broadsheet/tabloid spectrum. The terms broadsheets and tabloids stem from paper format—the earlier being printed on A1 paper and the later in variations of the smaller A3 paper—but have been adopted to refer to editorial decisions that define newspapers' journalism standards, setting quality press (broadsheets) apart from popular press (tabloids). We expect the analysis of the data to indicate whether social media users interact with content in a way that validates the division between broadsheet and tabloids perceived in print newspapers.

Broadsheet papers use investigative approaches to news that emphasize in-depth coverage and a sober tone in articles and editorials, with smaller headlines and fewer pictures that are expected in high-quality journalism (Preston, 2004). While broadsheet newspapers emphasize hard news coverage, fact-checking and research based on a timeline in which the story unfolds, tabloid newspapers present on average lesser detailed articles, often directed by marketing departments and heavily influenced by demographic appeal and audience share. According to Andersen (2003), tabloid readership is on average younger and less educated. The detailed coverage of political issues offered by broadsheets also results in a readership more interested in politics and more likely to be well informed.

The rationale of this study is to assess the editorial profile of online newspapers in the face of social media feedback to news items. The editorial profile of newspapers is defined by the emphasis news outlets place on specific topics. The editorial profile of newspapers *Wall Street Journal, Financial Times, Financial Times Deutschland*, and *Valor Econômico* are defined by a strong emphasis on business and economic news, while *Los Angeles Times*—likely due to the location of the newspaper's headquarters in the Los Angeles—presents a higher-than-average occurrence of infotainment news and a special emphasis on the American film industry. A similar prevalence of soft news content is found on newspapers *New York Post* and *New York Daily News*, while newspapers *The New York Times, The Guardian*, and *The Independent* present a higher-than-average emphasis on hard news topics, particularly Politics, Economy, and World news.

Data

Forty Newspapers Spanning Five Countries

We monitored tweets containing links to news articles from eight of the largest national newspapers in the United States, United Kingdom, Spain, Brazil, and Germany. The data span from Monday, October 1, 2012 (00:00:00 Greenwich Mean Time [GMT]) to Sunday, October 14, 2012 (23:59:59 GMT), thus including 2 weeks of data collection. News outlets were selected based on the circulation of print newspapers in Germany (Informationsgemeinschaft zur Feststellung der Verbreitung von Werbeträgern e.V., 2012), Spain (Oficina de Justificación de la Difusión [Circulation Audit Bureau; OJD], 2012), Brazil (Associação Nacional de Jornais, 2011), the United Kingdom (Audit Bureau of Circulations [ABC], 2012), and the United States (Alliance for Audited Media, 2012).

The data were retrieved using the software yourTwapperKeeper (O'Brien, 2010) that connects to Twitter streaming API and allows for keyword monitoring. The dataset includes a total of 2,842,699 tweets divided into 123,191 from German news outlets, 394,533 from Brazilian news outlets, 792,952 from Spanish news outlets, 537,606 from British news outlets, and 994,417 from American news outlets. From the 2,842,699 tweets containing links to news articles in the dataset, 64,605 are mention-messages (AT) and 1,026,296 are retweet-messages.

For the purposes of this investigation, we only analyzed national newspapers and excluded magazines and tabloid-journalism newspapers. Therefore, the dataset does not include British tabloids *Daily Mirror, Daily Mail*, and *The Sun*; German tabloid *Bild*; Spanish news outlets *Qué!*; American tabloids *The Globe and National Enquirer*; and Brazilian newspapers *Super Notícia, Extra*, and *Correio do Povo*. It does not include the magazines *The Economist, Der Spiegel, Focus, Veja, Newsweek*, and *The New Yorker*. British newspapers *The Sunday Telegraph* and *The Independent* on Sunday were not included because they lack a dedicated Internet website and depended on their sister papers to relay news articles on the Internet.

Although on average more readers read tabloids than broadsheets (Murphy, 2003), tabloid newspapers like *Bild, The Sun*, and *Enquirer* were not included in this study due to the editorial styles not conforming to the principles of quality press that we investigate in this article. Moreover, the significantly higher circulation of tabloids can skew the dataset and strongly affect the analysis of national newspapers. London-based newspaper *Daily Express*, though arguably a tabloid-newspaper, was selected to fill the position of the eighth national newspaper in the United Kingdom due to the circulation figures similar to national quality papers. Table 1 shows all news outlets investigated in this study.

Data Cleaning and Processing

We removed spam messages from the dataset and tabulated the data according to the attention share (sheer number of links) received by each news outlet and the overall volume of links per country. The processed data shows the circulation of highly retweeted and mentioned messages, the relative occurrence of overlapping users, and the concentration of links to news articles posted by top users. About one third of the dataset (35%) includes links to American news outlets, while Spanish newspapers account for more than a quarter (28%) and British newspapers for about a fifth (19%) of the dataset. The remaining fifth of the dataset includes Brazilian (14%) and German (4%) news outlets. Figure 1 shows the distribution of links to news outlets in the United States,

Table 1. Eight of the Largest Papers by Circulation in the United States, United Kingdom, Spain, Brazil, and Germany.

The United States	The United Kingdom	Spain	Brazil	Germany
The Wall Street Journal	*The Sunday Times*	*El País*	*Folha de S.Paulo*	*Die Zeit*
USA Today	*Daily Telegraph*	*El Mundo del Siglo XXI*	*O Estado de S.Paulo*	*Süddeutsche Zeitung*
The New York Times	*Daily Express*	*ABC*	*O Globo*	*Frankfurter Allgemeine*
Los Angeles Times	*The Times*	*La Vanguardia*	*Zero Hora*	*Die Welt*
New York Post	*Financial Times*	*El Periódico Catalunya*	*Estado de Minas*	*Frankfurter Rundschau*
Daily News	*Evening Standard*	*La Razón*	*Valor Econômico*	*Der Tagesspiegel*
The Washington Post	*The Guardian*	*El Correo*	*Correio Braziliense*	*Financial Times DE*
Chicago Sun-Times	*The Independent*	*La Voz de Galicia*	*Jornal do Commercio*	*die Tageszeitung*

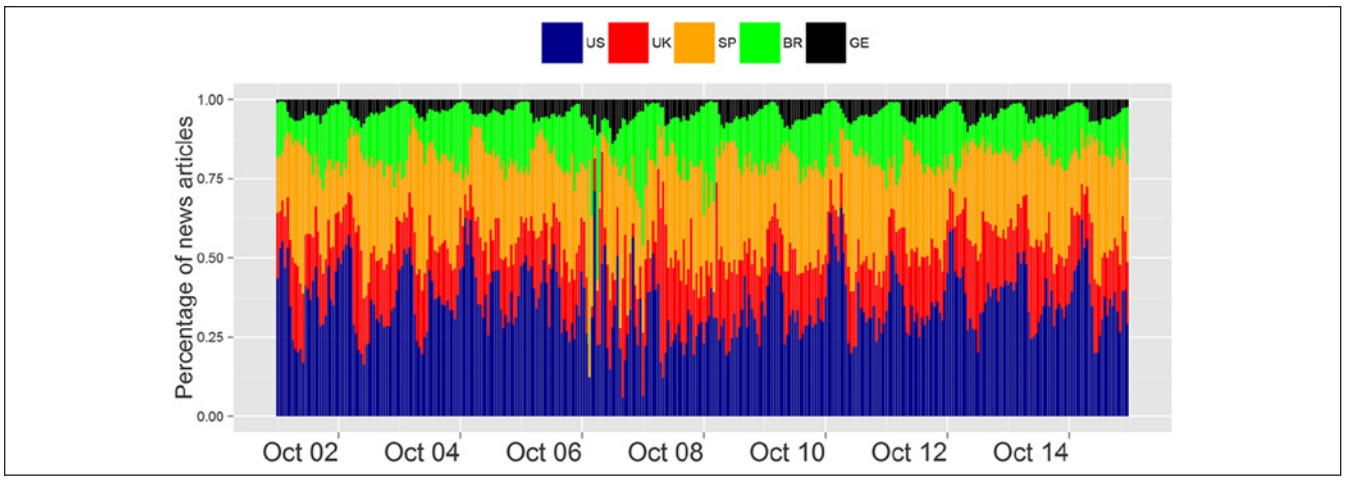

Figure 1. Percentage of links to news articles of American, British, Spanish, Brazilian, and German news outlets.

United Kingdom, Spain, Brazil, and Germany during the 14 days of data collection.

News Articles Timeline

Visualizing the longitudinal distribution of news articles through the day is challenging because of the multiple time zones in the United States and Brazil and the different time zones in the United Kingdom, Spain, and Germany. Time zones were adjusted to Coordinated Universal Time (UTC)+0 (GMT), using Eastern Standard Time (EST) zone as reference for news articles from the United States, Brasília for Brazil, and the official times for the United Kingdom, Spain, and Germany. Figure 2 shows the timeline of news articles through the day and indicates an abrupt peak of American links to news articles as early as 05:00 followed by subsequent peaks at 09:00, 14:30, and again at 17:30. Links to news articles in other countries follow a similar pattern, yet the peak is not as accentuated as observed in the United States. Links to news articles of Spanish newspapers present a slightly different pattern, increasing steadily from 06:00 to 11:00 and decreasing abruptly from 12:00 to 16:00, likely due to Spaniards routine of slowing down for lunch around 13:00 and the institution of midday sleep (*siesta*) from 14:00 to 16:00.

Twitter Clients

The processed data indicate which Twitter client was used to post links to news articles, thus allowing for the identification of news articles being actively distributed by users or by automatic applications. We found that twitterfeed, a utility that allows the feeding of content from website Rich Site Summary (RSS) feeds to Twitter; HootSuite, a social media management system for brand management; and dlvr.it, an application that automatically delivers content to paid media channels, were responsible on average for 25% of all links to news articles. The remaining 75% is mainly distributed across Twitter website (19%), Tweet Button (12%), Twitter for iPhone (9%), and Twitter for Android (7%). The data are consistent with previous reports (Sysomos, 2009) that found fewer than half of tweets being posted using the web user interface, and that most users used third-party applications. Figure 3 shows the distribution of Twitter clients per country.

Due to the high diversity of clients operating within the Twitter system—as of July 2011, Twitter's ecosystem comprised over one million applications, services, and clients (Twitter, 2011)—We manually coded Twitter services, clients, and applications into four groups according to user's usage: desktop, mobile, feeds, and managers. These groups

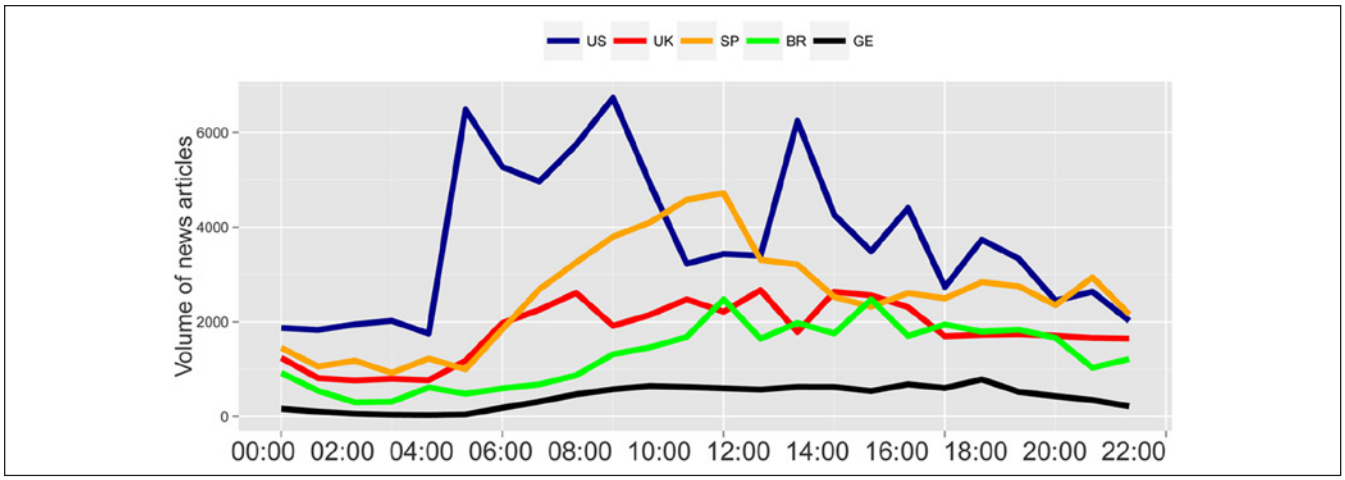

Figure 2. Timeline (24-hr) of news articles posted on Twitter in the United States, United Kingdom, Spain, Brazil, and Germany. Times zones adjusted to GMT (U.S. time zone reference: EST; BR time zone reference: Brasília).
Note. GMT = Greenwich Mean Time; EST = Eastern Standard Time; BR = Brazil.

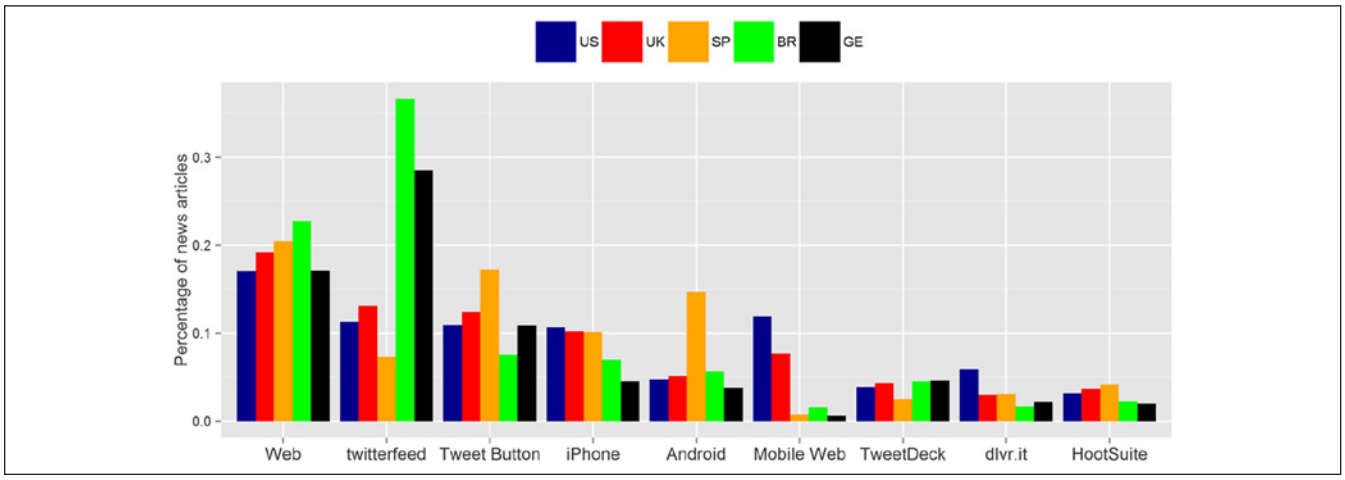

Figure 3. Distribution of Twitter clients used to post links to news articles per country.

indicate whether links were distributed by active users or generated by automatic updates. The first group includes desktop applications and web services that depend on user's input to feed Twitter's microblogging service via desktop software or Internet browsers. The second group also requires user's input and includes news articles posted using mobile platforms, including tablets and mobile phones.

The third group contains news items relayed automatically by bots or news feed, so that content is frequently updated and distributed on Twitter from the publisher's website. The last group refers to social media management platforms, which integrate a number of social networks and provide elementary analytics. Social media optimization (SMO) services that aim at increasing the awareness of a product, brand, or event, such as SocialFlow and dlvr.it, were split between the third (feeds) and the fourth (managers) groups depending on the company's business model.

Figure 4 shows the distribution of Twitter clients used to post news articles across countries and the distribution of Twitter clients per country. Even though Desktop (39%), Mobile (34%), and Feeds (25%) are the predominant source of news articles across the dataset, Figure 4 shows a significant variance across countries. Countries with a large number of highly active users, such as the United States, United Kingdom, and Spain present a much higher percentage of links to news articles from mobile platforms, while countries with fewer highly active users present a lower incidence of mobile platforms and a much higher percentage of items automatically relayed by news aggregators (news feeds) and bots.

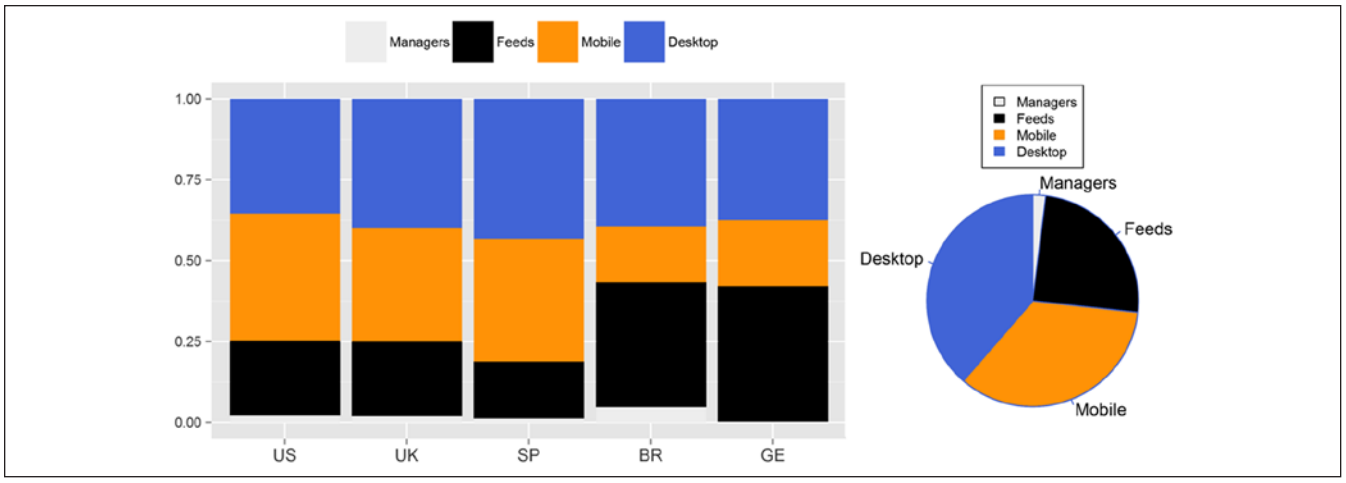

Figure 4. Twitter clients used to release news article per country (left) and overall ratio of Twitter clients (right).

Method

Singletons

We analyzed the data to understand how news items are released on Twitter and the prevailing news sections in each country. We relied on the classification of Twitter clients to investigate the relationship between the diversity of Twitter clients and the replication of news articles on Twitter. We identified the news outlets that resorted to Twitter as a mere broadcasting channel and the outlets that relied on Twitter to engage with their audiences by calculating the percentage of messages with links to news article that did not generate any conversation by means of RT, AT, or replication of the link to the news article. We called these messages singletons, as they refer to a set of messages with exactly one element (replication = 0). We compared the percentage of singletons with the percentage of messages relayed automatically by bots or news aggregators (feeds).

Resolving Twitter Links

Links shared on Twitter are automatically shortened to the http://t.co wrapper, and as a result we had to resolve 2,842,699 links before classifying news articles into news sections. Twitter links were expanded using a three-pass routine to resolve nested shortened URLs (i.e., previously shortened bit.ly URLs that are again shortened by the t.co URL wrapper) and to retrieve the target URLs. After resolving the addresses, we removed links that returned error, spam, or sweepstakes (promotional offers).

Hard-paywalled links allow minimal to no access to content without paid subscription, and most of the links pointed to British newspapers *The Times* and *The Sunday Times*, and to American news outlet *The Wall Street Journal*. Articles from news outlets that have adopted "hard paywalls" amounted to 20,655 news links that could not be classified and were thus removed from the dataset. We also removed links that failed to be replicated or that pointed to newspapers' Picture of the Day, as the varied content could not be classified into any newspaper section. Tables 2 and 3 show the process of converting Twitter raw data with tweet and parsed shortened URL into a table with resolved URLs and the section of the news articles.

Assigning News Sections

News articles with web URL link format that did not include the news section of articles impose further challenges to the analysis. American newspaper *The Wall Street Journal*, German newspaper *die Tageszeitung*, British *The Sunday Times*, and Spanish daily newspapers *El Correo, El Periodico*, and *La Razón* adopted domain structures that frequently lack the section of the news article. In these cases, we performed a customized search with regular expression matching to identify the newspaper sections. Blogs affiliated with media outlets also posed challenges to the classification and required customized regular expressions given the wide variety of domain structure observed, but the final processed data also include blog posts from journalists affiliated to the investigated newspapers.

El País popular special section "Viñetas," commentator articles, and open-editorials were grouped in the section Opinion, which also includes editorials and newspapers' columnist articles. A number of German newspapers have a special section on bicycles, particularly Die Zeit, and these news articles were grouped in the section Cars because they refer to bicycles as a means of transport. Articles on parenting were classified in the Lifestyle section, elections in the Politics group, obituaries in Local News, Career in Jobs, Travel advice in Tourism, and Videogames in Entertainment (not in Technology). Table 4 shows a column with news sections across countries and the most common labeled news

Table 2. Twitter Raw Data With Parsed Shortened URL, User, and Tweet-Message.

Original URL	User	Original Tweet
http://t.co/dli5BWxs	Userxxx1	RT @govewatch: Good pic of Michael Gove literally crushing one child's aspirations in person. (via http://t.co/dli5BWxs) #GoveMUSTgo http://t.co/8ECos4Nq
http://t.co/69mKflyb	Userxxx2	RT @n64k: Italy kills tax-deduction for churches! Let's do it HERE. Why let taxes support delusional hatred orgs? http://t.co/69mKflyb @RichardDawkins
http://t.co/K1Sb1ysE	Userxxx3	RT @alexispetridis: I confess I'm slightly amazed at how beautiful and moving Peaches Geldof's article on gay marriage is: http://t.co/K1Sb1ysE
http://t.co/SsdU51zg	Userxxx4	David Cameron spoke of defending disabled people. It is beyond shameless. This is what his government is really doing http://t.co/SsdU51zg
http://t.co/clp6pWrb	Userxxx5	RT @GuardianUS: Fun fact: Paul Ryan was three years old when Joe Biden was elected to Congress. True! http://t.co/clp6pWrb via @RichardA
http://t.co/9vhvqmjL	Userxxx6	RT @12RAED: @NabilAlawadhy الاندبنـــــدنت :الحمل حالات من ٪٥٤و ،خلقيـــا مشـوهين الفلوجـــة مواليـد نصـف من أكـثر والإجهاض الجنيـــن بوفـــاة تنتهـــي http://t.co/9vhvqmjL. . .

Note. RT = retweet messages.

Table 3. Twitter Processed Data With Shortened URL, Resolved URL, Newspaper, and News Section.

Original URL	Resolved URL	Newspaper	Sections
http://t.co/dli5BWxs	http://www.standard.co.uk/news/politics/tory-mp-slams-michael-gove-over-pace-of-education-reforms-8194612.html	Standard	Politics
http://t.co/69mKflyb	http://www.telegraph.co.uk/finance/financialcrisis/9598148/Italian-church-to-be-stripped-of-tax-exemption-from-2013.html	Telegraph	Economy
http://t.co/K1Sb1ysE	http://www.independent.co.uk/voices/comment/peaches-geldof-joins-the-independent-voices-campaign-to-legalise-samesex-marriage-in-britain-8204345.html	Independent	Opinion
http://t.co/SsdU51zg	http://www.independent.co.uk/voices/commentators/owen-jones-david-cameron-praises-paralympians-but-his-policies-will-crush-them-8082036.html	Independent	Opinion
http://t.co/clp6pWrb	http://www.guardian.co.uk/world/richard-adams-blog/2012/oct/11/vp-debate-ryan-biden-live	Guardian	World
http://t.co/9vhvqmjL	http://www.independent.co.uk/life-style/health-and-families/health-news/iraq-records-huge-rise-in-birth-defects-8210444.html	Independent	Lifestyle

sections in each country (sample from a bag-of-words with 1,616 words).

The final dataset includes 693,066 news articles each classified into one of the following 21 news sections: Arts, Cars, Crime, Economy, Education, Entertainment, Environment, Fashion, Health, Jobs, Lifestyle, Local news, National news, World news, Opinion, Politics, Science, Sports, Technology, Tourism, and Weather. Our method of classification successfully resolved, identified, and assigned news sections to 693,066 news articles from a total of 1,013,286 URL links, thus successfully classifying 68% of the news articles. The processed data offers an interesting and novel overview of news articles that circulated on the Twittersphere during the period under investigation.

Clustering Analysis

We exploited our prior knowledge of the properties of the dataset and used the "elbow" method (percentage of variance explained as a function of the number of clusters) to determine the appropriate value of k. We found the optimal number of clusters to be 4 over 100,000 iterations on the relationship between newspapers, and 2 clusters over 100,000 iterations on the relationship between news sections. We applied a standard clustering algorithm (k-means) using Euclidean distance on the frequency of news sections per newspaper. We also calculated the hierarchic clustering by transforming the affiliation matrix into a distance matrix between newspapers and news sections. The first clustering analysis was performed using a matrix of pairwise dissimilarities (1-similarity) between newspapers, and the second clustering analysis was performed on a matrix of pairwise dissimilarities between news sections.

The aim of the clustering analysis is to evaluate whether user's preference for news sections and news outlets reflects the perceived editorial decisions made by newspapers and newspersons that separate newspapers into broadsheets and tabloids—the first group including news outlets with focus on hard news, and the later including newspapers that focus on infotainment and soft news events. The second clustering analysis on a matrix of news sections was performed to evaluate whether the online distribution of news articles is

Table 4. Combined News Sections Across Countries (First Column) and the Most Frequent Names for Sections Per Country.

Sections	The United States	The United Kingdom	Spain	Brazil	Germany
Arts	Arts	Arts	Cultura	Ilustrada	Feuilleton
Cars	Motor	Motoring	Motor	Veículos	Auto
Classified	Classified	Classified	Clasificados	Classificados	Anzeige
Crime	Investigations	U.K. crime	Justicia	Policia	Gewalt
Economy	Business	Economy	Economía	Dinheiro	Wirtschaft
Education	Education	Education	Educacion	Educação	Studium
Entertainment	Entertainment	Features	Entretenimiento	Acontece	Panorama
Environment	Environment	Sustainable	Medioambiente	Meio-ambiente	Umwelt
Fashion	Fashion	Beauty	Smoda	Moda	Mode
Health	Health	Health	Salud	Saúde	Gesundheit
Jobs	Jobs	Jobs	Empleo	Emprego	Beruf
Lifestyle	Style	Lifeandstyle	Estilo	Estilo	Lebensart
Local news	Metro	London	Ciudades	Cotidiano	Gesellschaft
National	Country	Society	Nacional	Brasil	Deutschland
Opinion	Editorials	Comment	Opinión	Opinião	Kolumnen
Politics	Washington	Westminster	Política	Poder	Politik
Science	Science	Science	Ciencia	Ciência	Wissen
Sports	Sports	Sports	Deportes	Esporte	Sport
Technology	Technology	Technology	Tecnología	Informática	Digital
Tourism	Tourism	Travel	Elviajero	Viagem	Reisen
Weather	Weather	Weather	El Tiempo	Tempo	Wetter
World	World	International	Internacional	Mundo	Welt

consistent with traditional journalism classification of news items between hard and soft news. The results of these analyses are presented in the next section.

Results

Readership

We found that the number of links to news articles from AT varies little across news outlets of different countries ($\sigma X = 0.3\%$), but the number of links to news articles from retweet-messages (RT) varies considerably ($\sigma X = 7\%$), being significantly higher in Spanish news outlets (44%) and significantly lower in German newspapers (24%). We also found that the proportion of retweeted news articles in American, British, and Brazilian news outlets is fairly equal ($\sigma X = 1\%$). Figure 5 shows the number of links to news articles from tweets, RT, and AT for each of the 40 news outlets investigated in this study.

We found statistical significant correlations between the volume of tweets, retweets, and the number of connections between users that relayed a news article link. Leading newspapers within their respective countries presented statistically significant correlations between retweets and tweets ($r = .92$, $p < .001$) and between retweets and network topology ($r = .92$ and $r = .89$, respectively, $p < .001$), particularly the news outlets *The New York Times* and *Washington Post* in the United States, *The Guardian* and *Daily Telegraph* in the United Kingdom, *El País* and *El Mundo* in Spain, *Estado de S.Paulo* and *Folha de S.Paulo* in Brazil, and *Die Welt* and *Frankfurter Allgemeine Zeitung* in Germany.

We examined the concentration of links to news articles posted by top users and found that on average 5% of users are responsible for 50% of links to news articles across the five countries, being as high as 62% in Brazil and as low as 46% in the United Kingdom. Table 5 shows that on average 50% of all URL links to news articles are posted by just 5% of users, but our results are at odds with the importance of top users found in previous studies on the distribution of URL links. Wu, Hofman, Mason, and Watts (2011) reported that less than 0.05% of Twitter users attracted almost 50% of all attention within Twitter, while our results indicate that 0.05% of Twitter users posted on average only 10% of URL links to news articles and significantly less in highly active countries like the United States and the United Kingdom. The results also present a point of departure from previous investigations that identified the importance of gatekeeping in online newspapers (Dimitrova et al., 2003) and confirm the influence of a large community of least-active users (Bastos, Raimundo, & Travitzki, 2013) that posted one fifth of the links to news articles that circulated on Twitter in the period.

Table 6 shows the overlapping of users (shared readership) across different countries and within the same country. These figures refer to users that posted links to more than one news outlet within the same country, in the

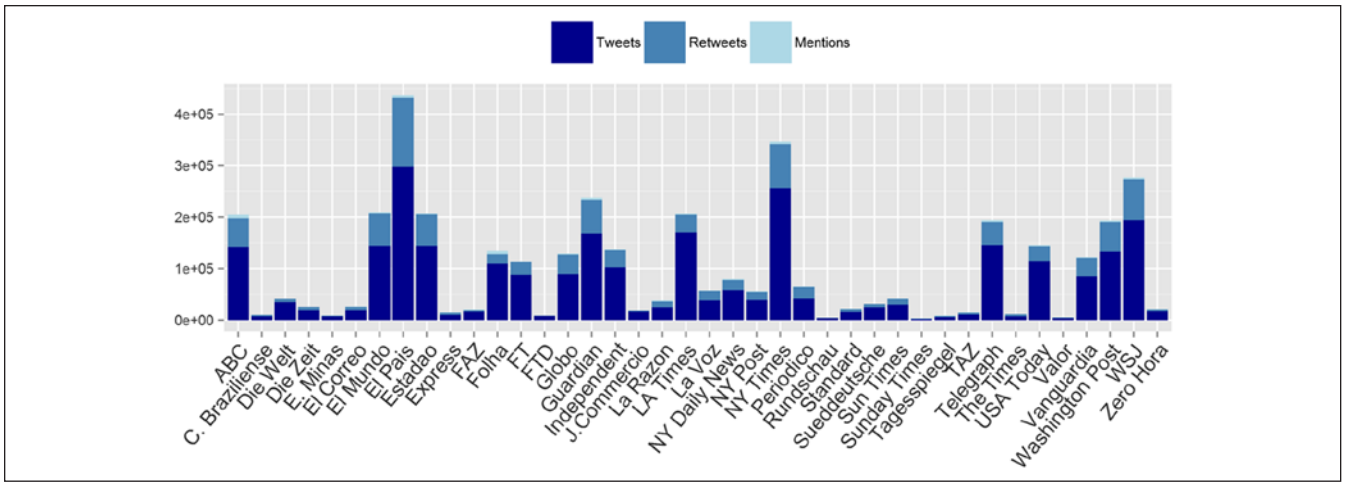

Figure 5. Number of links to news articles from tweets, retweets and mention-messages per news outlet.

Table 5. Participation of Top Users in the Distribution of Links to News Articles Per Country.

	The United States	The United Kingdom	Spain	Brazil	Germany	Total
Links to articles	994,417	537,606	792,952	394,533	123,191	2,842,699
0.05% top users	8%	8%	9%	17%	9%	9%
5% top users	47%	46%	46%	62%	61%	50%
10% top users	56%	55%	57%	69%	70%	59%
30% top users	73%	72%	75%	83%	84%	75%
50% top users	81%	80%	84%	89%	90%	83%

Table 6. Number of Unique Users (Readers) That Posted Links to More Than One Newspaper Within the Same Country (Column on the Left) and Number of Unique Users That Posted Links to Newspapers of Another Countries (Right Column).

Country	Cross-readership inside the country			Readership across different countries		
	Users	Overlapping users	Percent	Users	Overlapping users	Percent
The United States	422,808	81,638	19	422,808	53,568	13
The United Kingdom	256,275	43,701	17	256,275	52,233	20
Spain	372,726	104,199	28	372,726	21,353	6
Brazil	110,185	22,217	20	110,185	4,257	4
Germany	32,443	11,394	35	32,443	4,167	13

latter case, and to news outlets of different countries, in the former case. The table shows high cross-readership within Germany and Spain and low cross-readership within the United States, United Kingdom, and Brazil. These figures refer to the percentage of users that read and post links to more than one news outlet in the country and indicate that German and Spanish users read on average more different newspapers than their counterparts in Brazil, United Kingdom, and United States.

The table also shows that British users read a higher-than-average volume of news articles from foreign countries, likely due to the remarkably high cross-readership between American and British readers (40,368 users distributed links to American and British newspapers). The table also shows that Spanish and Brazilian readers are the least likely to read and distribute links to news articles of foreign countries. Perhaps

surprisingly, British readers are more likely to read and distribute content to supplementary newspapers outside the country than read and distribute links to additional British newspapers.

We compared the percentage of links to news articles per newspaper on Twitter with the newspapers' market share in print circulation, which was calculated based solely on the set of eight newspapers in each country. Table 7 shows the countrywise differences between the percentage of printed newspapers' circulation and the percentage of news articles on Twitter. We highlighted in red the figures of news outlets

Table 7. Circulation of Print Newspapers Compared With the Circulation of News Articles on Twitter.

The United States	Δ(%)	The United Kingdom	Δ (%)	Spain	Δ (%)	Brazil	Δ (%)	Germany	Δ (%)
WSJ	−6	*Sunday Times*	−27	*El País*	13	*Folha de S.Paulo*	5	*Die Zeit*	−9
USA Today	−11	*Daily Telegraph*	9	*El Mundo*	−1	*Estado de S.Paulo*	15	*Süddeutsche*	−1
NY Times	6	*Daily Express*	−16	*ABC*	4	*O Globo*	2	*FAZ*	−5
LA Times	10	*The Times*	−10	*La Vanguardia*	−1	*Zero Hora*	−11	*Die Welt*	15
NY Post	−3	*Financial Times*	6	*El Periódico*	−4	*Estado de Minas*	−5	*Rundschau*	−5
Daily News	−1	*Evening Standard*	−4	*La Razón*	−5	*Valor Econômico*	−4	*Der Tagesspiegel*	−2
W. Post	7	*Guardian*	25	*El Correo*	−4	*C. Braziliense*	−3	*Financial Times*	1
Sun-Times	−2	*The Independent*	17	*La Voz*	−1	*J. do Commercio*	1	*Die Tageszeitung*	6

Note. WSJ = *The Wall Street Journal*; FAZ = *Frankfurter Allgemeine Zeitung*.

experiencing a decline in the circulation of news links on Twitter relative to the figures of print newspapers. Sixty percent of the newspapers investigated in this study are experiencing a decline in the readership share in comparison with their print circulation.

USA Today, Daily Express, Zero Hora, and *Die Zeit* are the news outlets that registered the highest losses, while *Los Angeles Times, The Guardian, The Independent, El País, Estado de S.Paulo*, and *Die Welt* present a higher share in the volume of links to news articles on Twitter relative to the circulation of their printed newspapers. In Spain, even though the fluctuation is lower-than-average compared with other countries, we found that only *El País* and ABC news outlets presented an increase in the ratio between the circulation of printed newspapers and the diffusion of news articles on Twitter.

Broadcasting Singletons

News outlets with high levels of engagement on Twittersphere, such as *The Times, El País, El Mundo, Estado de S.Paulo, O Globo, Washington Post*, and *Wall Street Journal* consistently reported lower number of singletons and lower percentage of messages relayed by automatic aggregators (feeds). Based on this finding, we tested the hypothesis that news outlets that have most of their content posted via automatic feeds, using tools that link content from website RSS feeds to Twitter, are bound to generate low levels of interaction regardless of the volume of messages posted on Twitter. The results confirmed the hypothesis, showing a two-tier structure of highly interactive news outlets, with lower than average number of singletons and low number of tweets posted via RSS feeds, and a group of highly active, but not highly interactive group of news outlets, which resort to Twitter as another channel to rely content irrespective to feedback from users.

We found a significant correlation between news articles relayed by automatic feeds and news articles that failed to engage the audiences at the national level and with respect to each news outlet. We performed a statistical correlation analysis on the two variables and found that the proportion of singleton have a significant correlation with the proportion of tweets relayed by automatic feeds ($r = .78$, $p < .001$). We have not found any further positive significant correlation with the remaining Twitter clients used to post messages. Mobile and Desktop platforms presented negative correlations with the number of singletons ($r = –.64$ and $r = –.56$, respectively, $p < .001$), and social media management platforms did not present significant results ($r = .24$, $p < .001$). We conducted a *t*-test of statistical significance to compare the percentage of tweets posted from news feeds and the percentage of singletons on each news outlet. The results of the *t*-test within 40 newspapers rejected the null hypothesis of unequal variances, indicating symmetrical variances and positive correlation between tweets posted from news feeds and singletons—that is, tweets that did not generate any conversation, $t(39) = 31.82$, $p < .001$.

Based on these results, we fitted a linear regression model between the percentage of tweets posted from news feeds and the percentage of singletons on each news outlet. Even without removing outliers, the results of the linear regression reported a Multiple *R* of .78 with *R* Square of .61 and Adjusted *R* Square of .60 ($p < .001$), thus indicating that news articles released on Twitter by automatic feeds tend to generate little to no reaction and are unlikely to be read. Figure 6 shows the percentage of singletons and the percentage of tweets posted from news feeds with the fitted linear regression model. Figure 7 shows the percentage of replicated news articles and the percentage of tweets posted from desktop Twitter clients with the fitted linear regression model.

The results indicate a linear relationship between news articles posted from news feeds and news articles that are not replicated. Conversely, there is a linear relationship between news articles posted from desktop Twitter clients and messages that are replicated. The results indicate that the news outlets' strategy of automatically broadcasting their content to social media is not effective, as the links to news articles are not replicated and likely remain ignored. On the other hand, news articles posted from Twitter desktop clients present a linear relationship with message replication, therefore suggesting that news articles actively posted and commented

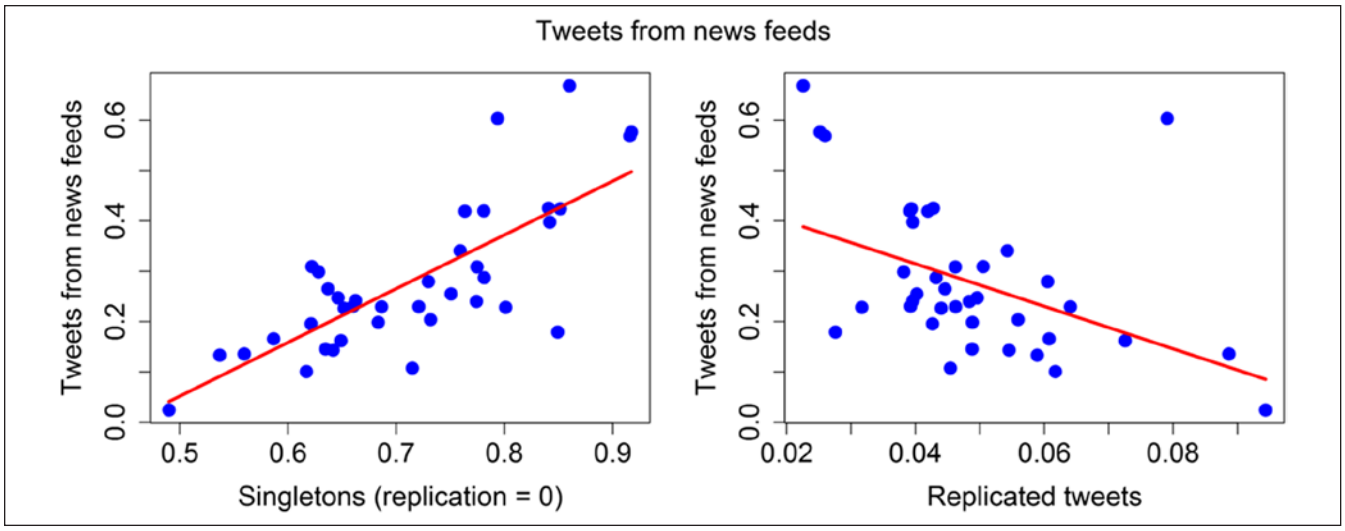

Figure 6. Tweets posted from news feeds with percentage of nonreplicated and replicated tweets.

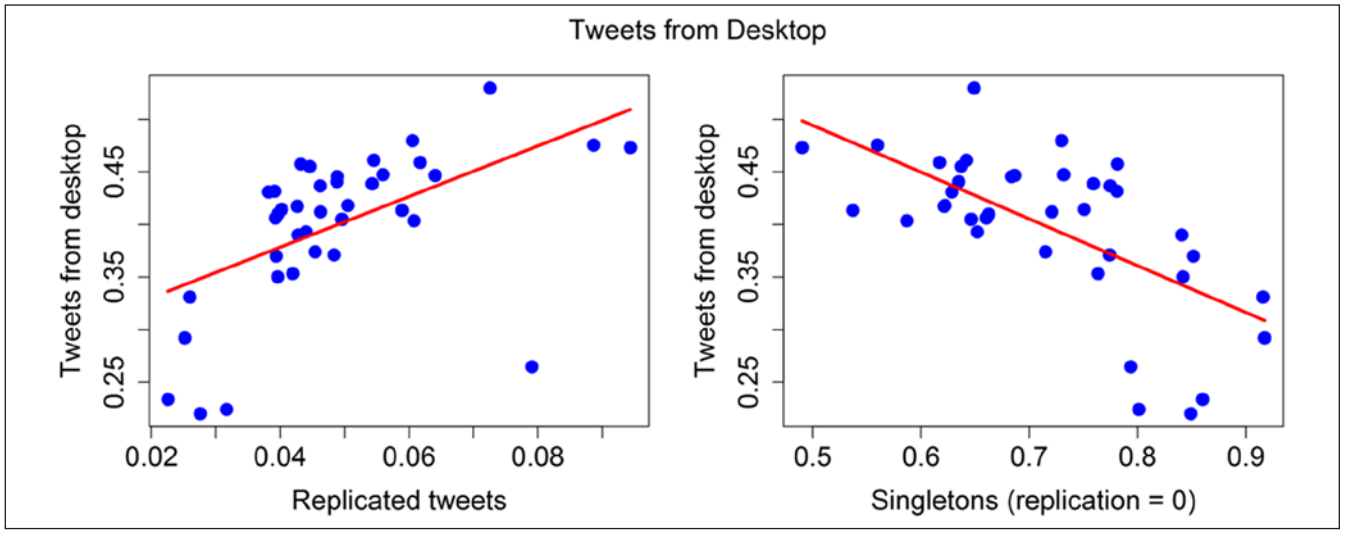

Figure 7. Tweets posted from desktops with percentage of replicated and nonreplicated tweets.

on by users are more likely to be read again and replicated by other users.

News Sections Per Country

We found that the prominence of news sections varies considerably from country to country. American newspapers present a lower-than-average volume of articles about Arts, Economy, and Local news, and a higher-than-average volume of news articles covering Entertainment, Fashion, Opinion, National, and World news. Links to news articles from British newspapers are more often Opinion, Environment, Education, Lifestyle, and World news pieces, and less often about Politics, Local, and National news. Spanish news articles present a higher-than-average incidence of Local and National news and a lower-than-average incidence of news articles on Politics and Sports.

Links to Brazilian news articles proportionally emphasize Sports, Arts, Education, Technology, and National news, but also present a lower-than-average volume of news articles about Economy, Opinion, and World news. German news outlets present a much higher-than-average volume of news articles about Economy and Politics and much lower-than-average volume of news articles about Entertainment, Opinion, and National news. Figures 8 shows the absolute number of news articles per section across countries and Figure 9 shows the normalized (relative) performance of each news section per country.

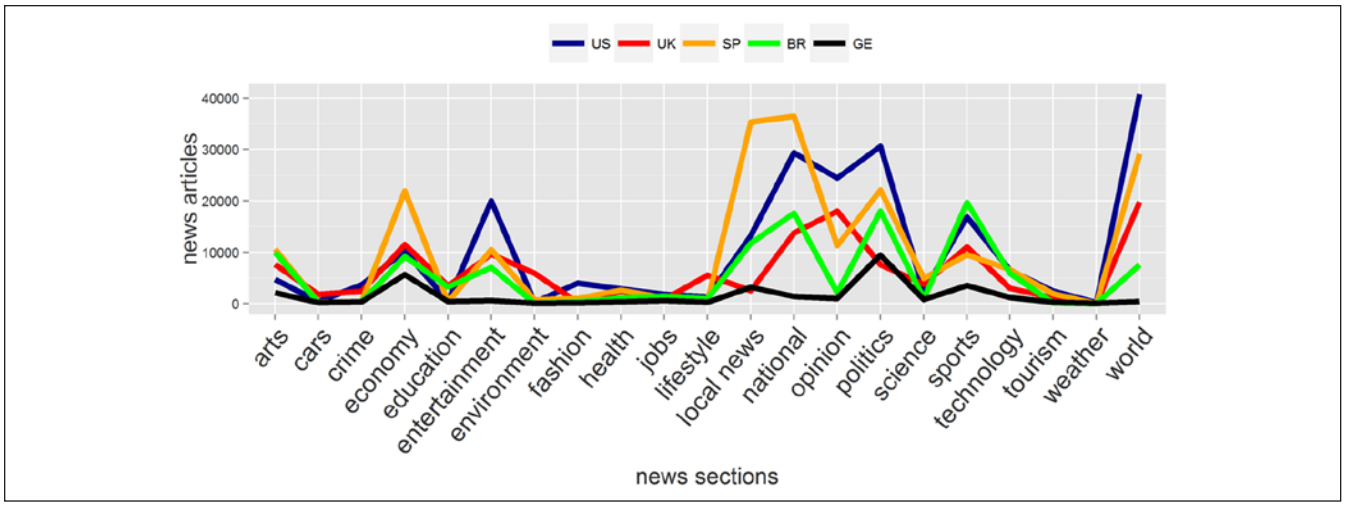

Figure 8. Countrywise distribution of news articles per newspaper section (absolute data).

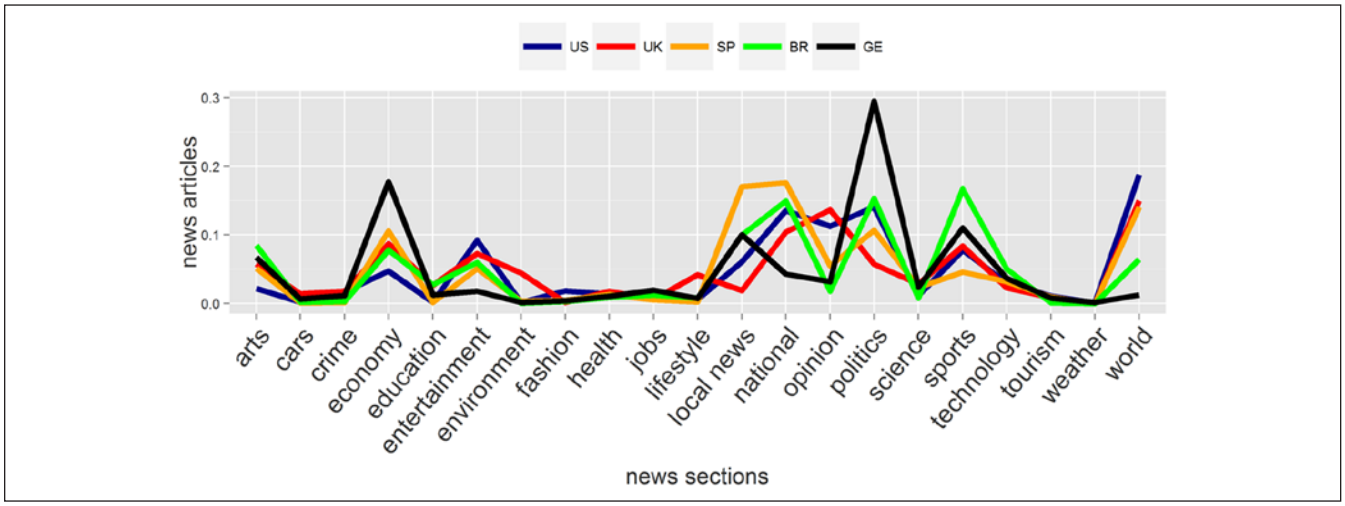

Figure 9. Countrywise distribution of news articles per newspaper section (normalized data).

Broadsheets Versus Tabloids

We investigated whether the separation between broadsheet and tabloids used in print newspapers is consistent with the content distributed online. Although we have focused on quality newspapers, the editorial styles of news outlets investigated in this study present a variation that encompasses the broadsheet and tabloid spectrum. We confirmed the hypothesis that newspapers can be grouped by cluster algorithms based on the relative volume of news articles covering hard and soft news topics, and the groups classified by the clustering analysis are consistent with the perceived editorial style of the news outlets.

The results revealed a clear division between news outlets that emphasize infotainment news and news outlets that cover hard news stories. The clustering analysis with *k*-means algorithm reported four clusters comprehending 11, 12, 5, and 12 news outlets, respectively. Group 1 clustered papers with a large volume of soft news items, including *Los Angeles Times, New York Daily News, New York Post, Chicago Sun-Times, USA Today, The Sunday Times, El Correo, La Vanguardia, Estado de S.Paulo, Jornal do Commercio*, and *Zero Hora*. Group 2 clustered most German newspapers and includes *The Washington Post, Evening Standard, El Periódico, Folha de S.Paulo, Valor Econômico, Frankfurter Allgemeine Zeitung, Frankfurter Rundschau, Financial Times Deutschland, Süddeutsche Zeitung, Die Tagesspiegel, Die Welt*, and *Die Zeit*.

Group 3 includes local news outlets or newspapers with a larger-than-average volume of local news, including *Daily Express, ABC, La Razón, La Voz de Galicia*, and *Correio Braziliense*. Group 4 includes the leading newspapers in most countries and also clustered news outlets with an editorial emphasis on high journalism standards. This last group

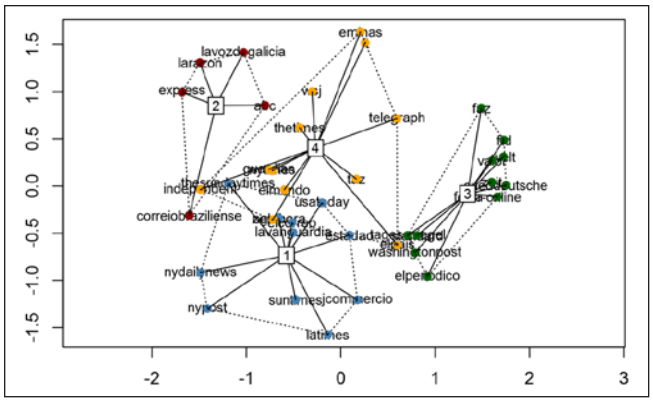

Figure 10. Newspapers clustered (dissimilarity measure based on Euclidean distance) into four groups according to the relative distribution of news articles per news section. Ordination axes indicate the total variation in the dataset.

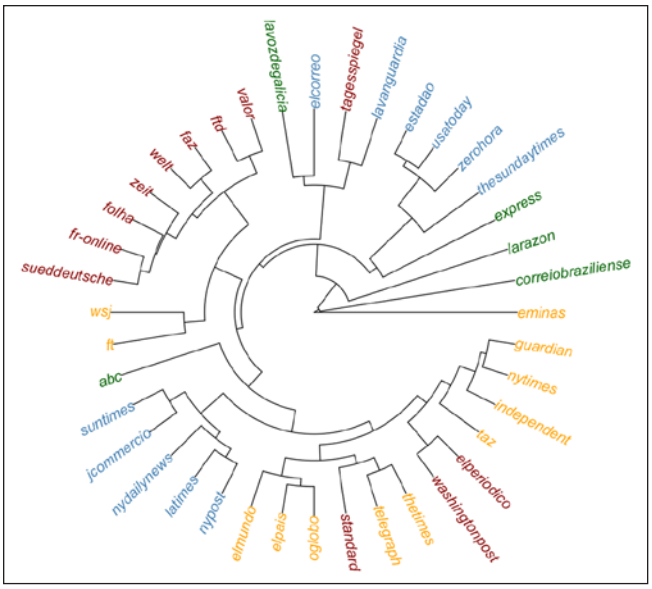

Figure 11. Newspapers clustered based on the relative distribution of news articles per section. Colors indicate *k*-means clustering, and the polar dendrogram is based on hierarchical clustering converted into a fan phylogram.

comprehends *The New York Times, The Wall Street Journal, The Guardian, The Independent, Daily Telegraph, Financial Times, The Times, El Mundo, El País, Estado de Minas, O Globo*, and *die Tageszeitung*. Except for German newspaper *die Tageszeitung* and Brazilian newspaper *Estado de Minas*, the news outlets clustered in Group 4 are leading newspapers within their respective countries and present a higher-than-average volume of hard news articles. Figure 10 shows the 40 newspapers clustered into four groups according to the relative distribution of news articles per news section.

The results from the hierarchical clustering based on squared Euclidean distances between the 40 news outlets returned a dendrogram that clustered newspapers with high volume of news articles dedicated to hard-news stories, such as Politics, Opinion, Local, National, and World news. The center of the dendrogram includes a much larger group of newspapers that also covered hard-news topics, but which also presented a higher volume of soft-news material. Figure 11 shows a circular dendrogram that clusters liberal news outlets *The Guardian* and *The New York Times*, and proceeds including newspapers with similar editorial profile, including *The Independent, Washington Post*, and *die Tageszeitung*. The subsequent clusters include leading British newspapers *The Times* and *Daily Telegraph*, followed by leading Spanish and Brazilian newspapers *El País* and *O Globo*.

Not surprisingly, Figure 11 shows business-oriented news outlets *Financial Times* and *The Wall Street Journal* clustered together (based on *k*-means and hierarchical clustering) on the left part of Figure 11, and also *Valor Econômico* and *Financial Times Deutschland* on the top of the circle. German leading newspapers *Die Zeit, Süddeutsche Zeitung*, and *Frankfurter Allgemeine Zeitung* are also clustered together on the top left of the ring. As expected, newspapers with a high volume of infotainment and soft news content clustered together at the lower-left portion of the circle, including news outlets *New York Post, Los Angeles Times, New York Daily News*, and *Chicago Sun-Times. Daily Express*, which as previously mentioned presents a sensibly different editorial style, but also newspapers with lower-than-average circulation like *Estado de Minas, Correio Braziliense*, and *La Razón*, fail to cluster with other newspapers and appear at the end of the circle.

Hard and Soft News

We also performed a clustering analysis based on squared Euclidean distances between the 21 news sections. We expected the relative distance between news sections to be consistent with the newsworthiness criteria that classify news articles into hard and soft news. The cluster analysis with *k*-means algorithm revealed a clear separation between hard and soft news topics without the need for any hand coding. The cluster analysis returned the two groups shown in Figure 12. The first group contains six sections and it clustered the hard news topics Economy, Opinion, Politics, and Local, National, and World news. The second group clustered 15 soft news sections, including Arts, Cars, Crime, Education, Entertainment, Environment, Fashion, Health, Jobs, Lifestyle, Science, Sports, Technology, Tourism, and Weather.

We also simulated the relationship between newspapers and news sections as a bipartite predator–prey network with two trophic levels (Figure 13). The lower level of the plot shows the most tweeted sections in increasing order of importance across countries, with Politics, National, World, and Economy as the most popular sections among Twitter

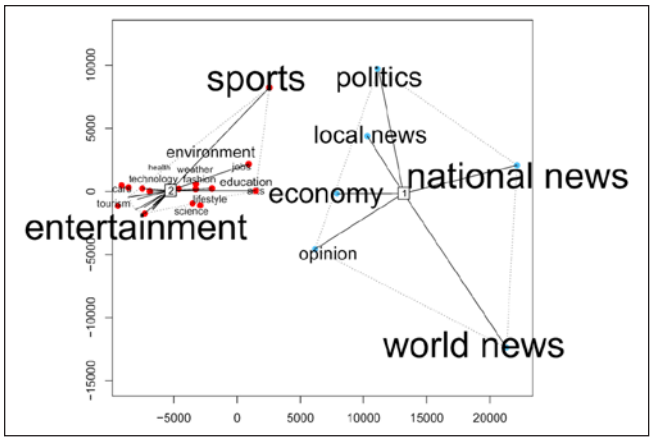

Figure 12. Clustering of news sections across newspapers with ordination axes indicating the total variation found in the dataset. Cluster 1 includes hard news stories while Cluster 2 groups soft news items.

users in the five countries, followed by Sports, Local news, Opinion pieces, Entertainment, and Arts. The upper level indicates the predominance of news sections per country. German newspapers present a higher-than-average volume of news articles focusing on Politics and Economy; Brazilian newspapers emphasize Sports and Politics; Spanish newspapers show a significant volume of articles on Local and National news, and American newspapers include a much higher-than-average volume of messages on Entertainment.

Figure 14 draws a food web as a grid using a matrix of news sections and newspapers. It shows three groups of newspapers. The group on the right includes leading newspapers within their respective countries, with a higher volume of hard news articles but also a considerable volume of articles across the range of news sections. The group on the left includes papers that emphasize soft news topics, and again we observe the same news outlets clustered with the *k*-means algorithm: *Los Angeles Times, New York Daily News, New York Post, Chicago Sun-Times, USA Today, The Sunday Times, El Correo, La Vanguardia, Estado de S.Paulo, Jornal do Commercio*, and *Zero Hora*. The group in the middle includes most of the German news outlets that presented a higher-than-average volume of news articles on Politics and Economy, but also an overall low volume of news items altogether.

Conclusion

The results reported in this paper indicate that the ecosystem of newsmaking is going through considerable changes. Audiences now have the opportunity to express their agency, not only as readers of texts but also as a fundamental piece that decides which news articles are replicated and which news sections gets the most attention across social networking sites. This investigation sheds light into the transformed architecture of journalism that increasingly relies on the participation of audiences (Mitchelstein & Boczkowski, 2009) and on the integration of multiple media ecosystems. We have shown how the distribution of news articles and sections change across countries and whether traditional models from print media endure in social networks.

The leading newspapers in their respective countries performed better than average on Twitter, particularly *El País* in Spain, *The New York Times* in the United States, *Estado de S.Paulo* in Brazil, *Süddeutsche Zeitung* in Germany, and *Daily Telegraph* in the United Kingdom. We established that the overall strategy of news outlets to relay links to news articles on Twitter via website feeds is inefficient, as most links remain singletons and likely unread. We also found that news articles tweeted using mobile and desktop platforms are much more likely to be replicated, and that countries with a larger volume of news articles also present a larger volume of mobile platform usage, though not necessarily a larger user base. These results confirm that feedback to news items from a large user base is pivotal for the replication of the content across social networks.

While most newspapers investigated in this study presented a decline in the readership share in comparison with their print circulation, the results also show that German and Spanish readers are more likely to read multiple national newspapers, and that British readers are more likely to resort to foreign sources of news than their counterparts in Brazil, United Kingdom, and United States. Contrary to previous research, we found the concentration of links to news articles posted by top users to be fairly low, as around a fifth of the links to news articles that circulated on Twitter in the period were posted by a large community of least-active users. The results also show that news readership on Twitter replicates the editorial profile of print publications, as new outlets are clearly grouped according to their editorial profile. The differences between broadsheet and tabloid inherited from legacy media remain relatively unchanged in terms of content shared on Twitter.

Consistent with previous research (Quandt, 2008), we found that different national cultures place different weights on news sections. Perhaps confirming and conflicting with the perceived stereotypes of nationalities reported in public opinion research (Prothro, 1954; Reigrotski & Anderson, 1959), German readers placed greater emphasis on Politics and Economy; Brazilians on Sports and Arts; Spaniards on Local and National news; Britons and Americans on Opinion and World news. On the other hand, and perhaps further accentuating positive and negative cultural stereotypes reported in social psychology literature (Cuddy et al., 2009; Jonas & Hewstone, 1986; Katz & Braly, 1933), Americans placed less emphasis on Economy and Arts; Britons on Politics and Local news; Spaniards on Sports and Politics; Brazilians on Opinion items and World news; and Germans on National and World news.

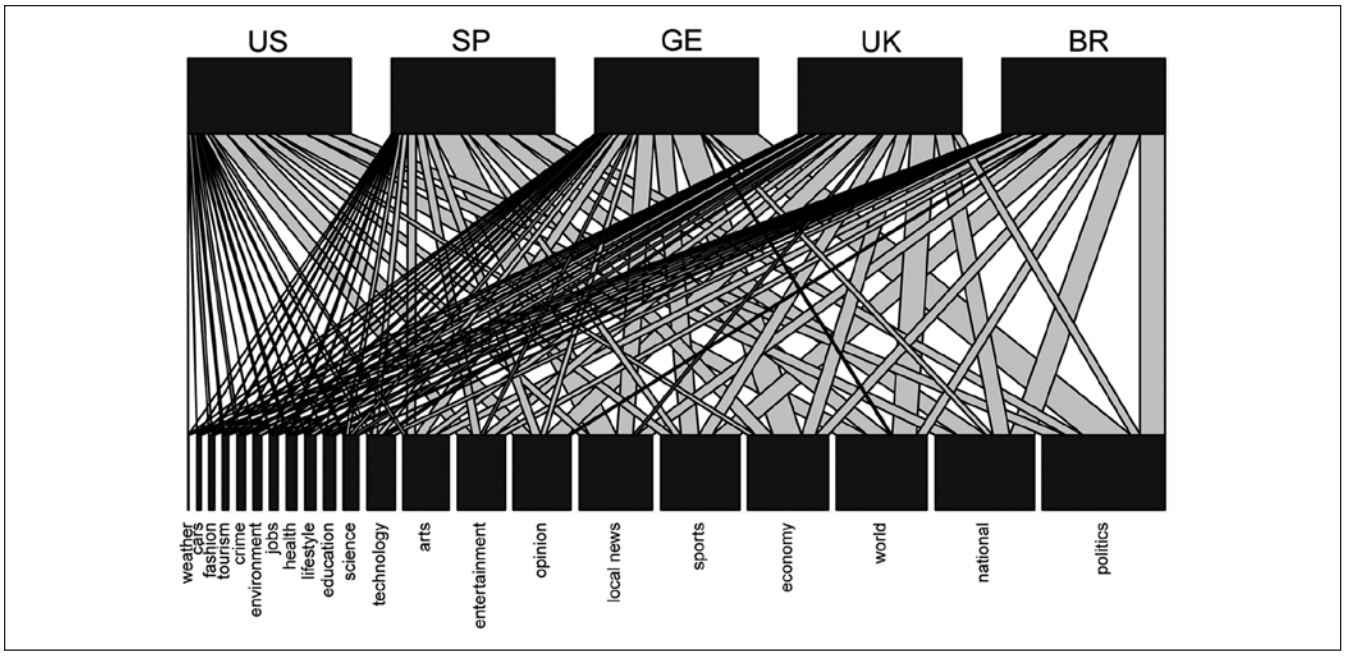

Figure 13. Most tweeted sections across countries (lower level) and prevalence of sections per country (upper level).

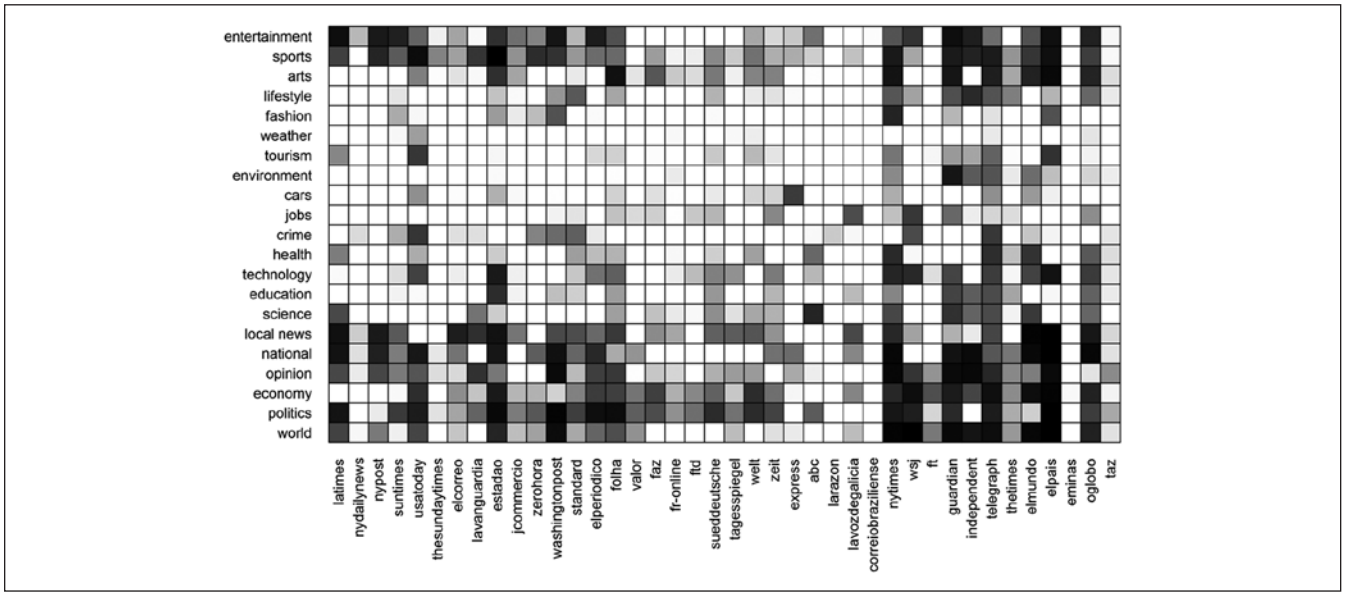

Figure 14. Frequency matrix of articles per news sections across the 40 newspaper. Data normalization based on squared Euclidean distances between the 21 news sections.

Finally, the methodology and the results presented in this paper are only possible due to the recent availability of social network data. Twitter in particular and social networks in general provide avenues for quantitative research at unprecedented scales. The understanding of which news sections of newspapers are the most and least read by readers in different parts of the world was until recently very restricted due to scanty data. Social media data not only allowed for this novel approach, but also likely changed the circulation and access to news items. We expect the availability of social network to data to support new lines of research inquiry and novel research methods designed to explore the networked architecture of online journalism.

Acknowledgments

The authors are thankful to Charlie Beckett for his insightful observations and suggestions. This research was completed while the first author was a visiting research fellow at the Department of Media and Communications of the London School of Economics and Political Science.

Declaration of Conflicting Interests

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

Funding

The author(s) disclosed receipt of the following financial support for the research and/or authorship of this article: This work was supported by the São Paulo Research Foundation (Grants 10/06243-4 and 11/22495-6).

References

Aitamurto, T., & Lewis, S. C. (2013). Open innovation in digital journalism: Examining the impact of Open APIs at four news organizations. *New Media & Society*, *15*, 314-331.

Alliance for Audited Media. (2012). *Top 25 U.S. newspapers for September 2012*. Arlington Heights, IL: Author.

Andersen, R. (2003). Do newspapers enlighten preferences? Personal ideology, party choice and the electoral cycle: The United Kingdom, 1992-1997. *Canadian Journal of Political Science(Revue canadienne de science politique)*, *36*, 601-619.

Associação Nacional de Jornais. (2011). *Maiores jornais do Brasil* [Largest newspapers in Brazil]. Brasília, Brazil: Author.

Audit Bureau of Circulations. (2012). *National newspapers report*. Berkhamsted, UK: Author.

Bastos, M. T. (in press). Shares, pins, and tweets: News readership from daily papers to social media.

Bastos, M. T., Raimundo, R., & Travitzki, R. (2013). Gatekeeping Twitter: Message diffusion in political hashtags. *Media, Culture & Society*, *35*, 260-270.

Bastos, M. T., Travitzki, R., & Puschmann, C. (2012, May). What sticks with whom? Twitter follower-followee networks and news classification. Paper presented at the Potential of Social Media Tools and Data for Journalists in the News Media Industry—Sixth International AAAI Conference on Weblogs and Social Media, Dublin, Ireland.

Conover, M. D., Ratkiewicz, J., Francisco, M., Goncalves, B., Menczer, F., & Flammini, A. (2011). *Political polarization on Twitter*. Bloomington, IN: Indiana University.

Cowan, M. (2010, November 22). Twitter co-founder hopes to create news network. *Reuters*. Retrieved from http://www.reuters.com/article/2010/11/22/us-twitter-news-idUSTRE6AL57Z20101122

Cuddy, A. J. C., Fiske, S. T., Kwan, V. S. Y., Glick, P., Demoulin, S., Leyens, J.-P., . . .Ziegler, R. (2009). Stereotype content model across cultures: Towards universal similarities and some differences. *British Journal of Social Psychology*, *48*, 1-33.

D'Haenens, L., Jankowski, N., & Heuvelman, A. (2004). News in online and print newspapers: Differences in reader consumption and recall. *New Media & Society*, *6*, 363-382.

Dimitrova, D. V., Connolly-Ahern, C., Williams, A. P., Kaid, L. L., & Reid, A. (2003). Hyperlinking as gatekeeping: Online newspaper coverage of the execution of an American terrorist. *Journalism Studies*, *4*, 401-414.

Eriksen, L. B., & Ihlstrom, C. (2000, January 4-7). Evolution of the Web news genre-the slow move beyond the print metaphor. Paper presented at the 33rd Hawaii International Conference on System Sciences, Manui, HI.

Fernandez-Rana, V., Simon, P. R., & Peserik, J. E. (1974). *USA Patent No. US3842719. 198/418.9*. Washington, DC: United States Patent and Trademark Office. Retrieved from http://patft.uspto.gov/netacgi/nph-Parser?Sect2=PTO1&Sect2=HITOFF&p=1&u=/netahtml/PTO/search-bool.html&r=1&f=G&l=50&d=PALL&RefSrch=yes&Query=PN/3842719

Filistrucchi, L., Fiesolana, B., & Domenico, S. (2005). *The impact of Internet on the market for daily newspapers in Italy*. San Domenico, Italy: European University Institute.

Franklin, B. (1997). *Newszak and news media*. London, England: Arnold.

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

Gentzkow, M. (2007). Valuing new goods in a model with complementarity: Online newspapers. *American Economic Review*, *97*, 713-744.

George, L. M. (2008). The Internet and the market for daily newspapers. *B. E. Journal of Economic Analysis & Policy*, *8*(1), 1-33.

Groth, O. (1928). *Die zeitung; ein system der zeitungskunde (Journalistik)* [The Newspaper: A system for the newspaper science]. Mannheim, Germany: J. Bensheimer.

Harcup, T., & O'Neill, D. (2001). What is news? Galtung and Ruge revisited. *Journalism Studies*, *2*, 276-280.

Hargittai, E., & Litt, E. (2011). The tweet smell of celebrity success: Explaining variation in Twitter adoption among a diverse group of young adults. *New Media & Society*, *13*, 824-842.

Hong, S. (2012). Online news on Twitter: Newspapers' social media adoption and their online readership. *Information Economics and Policy*, *24*, 69-74.

Informationsgemeinschaft zur Feststellung der Verbreitung von Werbeträgern e.V. (2012). *2012 Auflagenliste* [Circulation in 2012]. Berlin, Germany: Author.

Jonas, K., & Hewstone, M. (1986). The assessment of national stereotypes: A methodological study. *The Journal of Social Psychology*, *126*, 745-754.

Katz, D., & Braly, K. (1933). Racial stereotypes of one hundred college students. *The Journal of Abnormal and Social Psychology*, *28*, 280-290.

Kepplinger, H. M., & Ehmig, S. C. (2006). Predicting news decisions. An empirical test of the two-component theory of news selection. *Communications: The European Journal of Communication Research*, *31*, 25-43.

Larsson, A. O., & Moe, H. (2012). Studying political microblogging: Twitter users in the 2010 Swedish election campaign. *New Media & Society*, *14*, 729-747.

Messner, M., Linke, M., & Eford, A. (2011, April 1). *Shoveling tweets: An analysis of the microblogging engagement of traditional news organizations*. Paper presented at the International Symposium on Online Journalism Austin, Austin, TX.

Mitchelstein, E., & Boczkowski, P. J. (2009). Between tradition and change: A review of recent research on online news production. *Journalism*, *10*, 562-586.

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*, 1085-1102.

Murphy, G. L. (2003). The downside of categories. *Trends in Cognitive Sciences*, *7*, 513-514.

O'Brien, J., III. (2010). yourTwapperKeeper.

Oficina de Justificación de la Difusión. (2012). *Diarios (Daily Newspapers)*. Madrid, Spain: Circulation Audit Bureau.

Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O'Hara, R. B., & Wagner, H. (2012). vegan: Community Ecology Package. R package (Version 2.0–5).

Ovide, S. (2009, January 15). Twittering the USAirways plane crash. *The Wall Street Journal*. Retrieved from http://blogs.wsj.com/digits/2009/01/15/twittering-the-usairways-plane-crash/

Pérez-Peña, R. (2008, September 5). Times plans to combine sections of the paper. *The New York Times*, p. C3.

Pew Research Center. (2012). *Changing news landscape, Even television is vulnerable: Trends in News Consumption: 1991-2012*. Washington, DC: Author.

Preston, P. (2004). Tabloids: Only the beginning. *British Journalism Review*, *15*, 50-55.

Pritchard, D., & Hughes, K. D. (1997). Patterns of deviance in crime news. *Journal of Communication*, *47*(3), 49-67.

Prothro, E. T. (1954). Cross-cultural patterns of national stereotypes. *The Journal of Social Psychology*, *40*, 53-59.

Quandt, T. (2008). (No) news on the World Wide Web? *Journalism Studies*, *9*, 717-738.

R Development Core Team. (2009). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Reigrotski, E., & Anderson, N. (1959). National stereotypes and foreign contacts. *Public Opinion Quarterly*, *23*, 515-528.

Reuters Institute for the Study of Journalism. (2013). *Reuters Institute digital news report 2012: Tracking the Future of News*. Oxford, UK: Author.

Sankaranarayanan, J., Samet, H., Teitler, B., Lieberman, M., & Sperling, J. (2009, November 4-6). *TwitterStand: News in tweets*. Paper presented at the Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, Seattle, WA.

Schudson, M. (2011). *The sociology of news* (2nd ed.). New York, NY: W.W. Norton.

Schwarz, A. (2006). The theory of newsworthiness applied to Mexico's press. How the news factors influence foreign news coverage in a transitional country. *Communications: The European Journal of Communication Research*, *31*, 45-64.

Shoemaker, P., & Cohen, A. (2012). *News around the world: Content practitioners and the public*. New York, NY: Taylor & Francis.

Shoemaker, P., & Reese, S. (1996). *Mediating the message: Theories of influences on mass media content* (2nd ed.). White Plains, NY: Longman.

Sysomos. (2009). *Inside Twitter clients—An Analysis of 500 Million Tweets*. Toronto, Ontario, Canada: Author.

Tunstall, J. (1971). *Journalists at work: Specialist correspondents: Their news organizations, news sources, and competitor-colleagues*. London, England: Constable.

Twitter. (2011). *One million registered Twitter Apps*. San Francisco, CA: Author.

Wickham, H. (2009). *ggplot2: Elegant Graphics for Data Analysis*. New York, NY: Springer.

Wu, S., Hofman, J. M., Mason, W. A., & Watts, D. J. (2011, March 28-April 1). *Who says what to whom on Twitter*. Paper presented at the 20th international conference on World Wide Web, New York, NY.

Author Biographies

Marco Toledo Bastos is a postdoctoral fellow at the University of São Paulo, a lecturer and PhD researcher at the University of Frankfurt, and a visiting research fellow at the Department of Media and Communications of the London School of Economics and Political Science.

Gabriela Zago is a lecturer at the Federal University of Pelotas and a PhD candidate at the Communication and Information Graduate School of the Federal University of Rio Grande do Sul.