

The life and death of political news: Measuring the impact of the audience agenda using online data¹

Jonathan Bright, Oxford Internet Institute
Tom Nicholls, Oxford Internet Institute

Full article reference: Bright, J., and Nicholls, T. 2014. The Life and Death of Political News: Measuring the Impact of the Audience Agenda Using Online Data, *Social Science Computer Review*, 32(2), 170-181. Find it [here](#).

This is the accepted version of the article.

Abstract

The rapid development of online media as a major location for news consumption has stimulated a variety of debates about how journalism is changing in the internet era. Of particular importance have been worries about a potential turn towards populism, whereby journalists and editors shift away from reporting what is newsworthy to what their audience wants to hear, supported by the widespread availability of audience metrics. A wealth of ethnographic research has pointed to the potential importance of such statistics; but little quantitative work has been conducted to test for the existence of a relationship between audience behavior and editorial decisions.

This study seeks to fill that gap. Based on a novel dataset of over 40,000 articles published in 5 major UK news outlets over a period of 6 weeks, we explore the relationship between a news story's readership and its likelihood of being removed from the front page, based on the 'most read' lists common to many news websites. We find that being a 'most read' article decreased the short-term likelihood of being removed from the front page by around 25%, and that this effect was broadly similar for both

¹ We would like to acknowledge the helpful comments and suggestions made by Ralph Schroeder and John Laprise, and attendees at an Oxford Internet Institute seminar in June 2012. We would also like to thank three anonymous reviewers for their insightful critiques and suggestions for improvement. Any errors and omissions remain our own.

political and entertainment news. Surprisingly, we find a considerably greater influence in 'quality' publications than their tabloid counterparts. Our results are discussed as evidence of a still limited, but potentially developing, turn towards online populism.

Keywords: online journalism, audience metrics, digital media

Introduction

The news media play a crucial agenda-setting role in the political process. Through their ability to broadcast to large, loyal audiences in a sustained manner, journalists and especially editors have the potential, to an extent at least, to shape 'political reality' by deciding what is and is not important (McCombs, 2004). Hence the methods by which editors decide what stories to turn into news are a crucial area of research for scholars of political communication. Studies of the editorial process, which have a long history, have typically conceptualised it in terms of a series of gatekeepers (Shoemaker & Vos, 2009). Journalists, often working on distinct beats, choose from a pool of potential stories what to report to their editors; these editors then make a further selection, choosing which stories to publish, and where to publish them. It is these decisions which then go on to affect the broader political process. However, the developing importance of online news (Boczkowski, 2004, Paterson & Domingo, 2008) has the potential to alter this gatekeeping process in a variety of ways. In this article, we focus in particular on the changing role of the news audience in deciding what is news.

Whereas offline editors might have had only the vaguest idea of who their audience was and what they wanted, their online counterparts have detailed information about what articles are popular almost as soon as they are published (MacGregor, 2007,

Napoli, 2010, Anderson, 2011b). The importance of these metrics lies in the potential they have to alter the content of the news media. In particular, many authors (along with journalists and editors themselves) have expressed concern that they may be driving a culture whereby newsworthy but unexciting stories make way for more attention grabbing pieces, and journalistic values as a whole are undermined by a turn towards populism.

However, despite the potential significance of audience metrics, relatively little is known about their precise impact on the process of editorial decision making. In particular, while a variety of ethnographic and survey based studies have demonstrated that audience metrics are being incorporated into journalistic practice, hardly any research has tried to quantify what impact they have, if any (Lee, Lewis & Powers, 2012, p. 2); part of a broader tendency in online journalism studies to focus on the study of sites rather than individual stories (Karlsson & Strömbäck, 2010, pp.2-3).

This article seeks to remedy this deficit somewhat, by presenting a systematic, large scale study of the impact of user behaviour on the news media, with a focus on political articles. Exploiting the comparative ease with which online data can be collected, we examine the news cycles of five major UK news outlets across a period of six weeks. We capture the front page of each one of these websites every 15 minutes throughout this period, giving us a fine grained snapshot of how these news portals are evolving over time. On the basis of these data, we seek to explore the extent of audience participation in the political news cycle, looking at the extent to which readership statistics (measured through the proxy of the 'most read' sections of our sites) have an influence on the short-run likelihood of an article being removed from the front page.

The rest of the article is structured in the following way. We begin by presenting in more detail the importance of readership metrics for online news outlets, focussing on worries about the creation of a new online populism, and reviewing existing research on the subject. We move on to describing the characteristics of our dataset, showing how the different outlets under study have quite diverse publishing styles which are not easily reconciled with the traditional quality / tabloid division. Finally, we assess the impact of readership statistics on editorial decisions, arguing that strong viewing figures have the potential to reduce the short-term likelihood of an article being removed from the front page by 20-30%. This effect was common across political and entertainment news, and surprisingly much stronger in our ‘quality’ publications than their tabloid counterparts.

Online news: rise of the ‘audience agenda’?

The audience has always had an ambiguous role in the news production process (Loosen & Schmidt, 2012). While in theory one might expect journalists to be attentive to the needs of their listeners and viewers, research on the subject has tended to come to the opposite conclusion, characterising journalism as a profession with a “deliberate...ignorance of audience wants” (Anderson, 2011b, 553). While a variety of organisational and technical barriers made knowledge of the audience difficult in the offline era, this rejection of audience desires was (and still is) often portrayed as an important journalistic virtue, part of a more general set of news values which guide journalists and editors in their decisions about what constitutes news (Harcup & O’Neill, 2001). Journalism has sometimes been conceptualised as a type of public service (see Deuze, 2005, 447-448): and part of this service involves providing information which is

judged to have political relevance and importance, not just what audiences want to hear or find entertaining (MacGregor, 2007, 283).

The development of online journalism has prompted the revisiting of some of these previously settled conclusions about the role of the audience in news production (Bruns, 2008, Shoemaker & Vos, 2009, Napoli, 2010, Anderson, 2011a, Loosen & Schmidt 2012). Initially, much of this research was oriented towards the possibility of active audience participation in researching, editing and commenting on news stories, under various banners such as ‘user generated content’, ‘citizen journalism’ and ‘interactivity’ (see Oblak, 2005, Chung, 2007, Domingo, 2008). However more recently attention has been turning to what Anderson characterises as the potential influence of the ‘quantified’ audience (2011a). The online era has brought with it a wealth of new information about what readers of individual sites find important, allowing precise measurement of how many people visit a site, how often they come, what they choose to read when they arrive and how long they spend reading it (Deuze, 2003, 218, MacGregor, 2007, Domingo, 2008, 692, Shoemaker & Vos, 2009, 7). Furthermore, these statistics are frequently provided directly to writers, or displayed prominently in the newsroom itself (MacGregor, 2007, Anderson, 2011b). As online news consumption itself has grown in importance (Newman, Dutton & Blank, 2012, 9), the potential impact of these metrics is increasingly developing.

The rise of audience metrics has created concern both within the journalistic profession and academia, as part of a broader set of concerns about the way journalism is changing online (see e.g. Cohen, 2002, Brossard & Scheufele, 2013). In particular, the idea that news outlets will start to increasingly focus on stories which are likely to garner

traffic, rather than those which are inherently newsworthy, has been regarded as potentially problematic. As Shoemaker and Vos argue, “hard data about what readers want to read butts up against the social responsibility canon to give readers what they need to read” (Shoemaker & Vos 2009, 7). The incentives to follow the audience in this way are obvious: online business models rely on increasing traffic to a particular website, either in the hope of directing it to advertising websites or convincing it to pay for certain types of content. At a time when the business model of media outlets as a whole is under great strain, and many formerly profitable outlets are facing difficulty or closure, the impetus to ‘follow’ traffic may be considerable.

The consequences of such a shift for the broader agenda setting function of the news media could be significant: more prolific readers of news sites would play a disproportionate role in helping to select content; early readers of articles might have the ability to select out news which then never comes to the attention of viewers who read later in the day; and particular social classes or groupings which read news online less frequently might find their issues being subtly shifted down the agenda. Perhaps most importantly, the news as a whole could start to shift towards a more populist, ‘soft news’ style of publishing, where entertainment is prioritised over information.

This article seeks to assess the extent to which audience metrics influence editorial decisions in online news media. In particular, we focus on a new type of gatekeeping decision online editors have to take: the choice of when to remove an article from the front page by removing any link to it (or rather by replacing this link with one to a different article). While print editors had to decide what news to include in their paper, their online counterparts must also decide how long this news remains on the front page

of their site (Karlsson & Strömbäck 2010). Online articles themselves are rarely deleted: as the marginal cost of keeping content live is essentially negligible, and as even very old articles may occasionally attract search engine traffic or otherwise add value to the site (see Dezsö et al. 2006), pieces of news typically remain accessible through their web link indefinitely after their initial creation. However the existence of a link from the front page of the website, we assume, has major consequences both for article visibility and hence subsequent readership.

Our research question is simple: do readership statistics influence editorial decisions to keep articles on the front page? The extent to which such an influence exists has attracted little empirical research so far. A variety of ethnographic studies have demonstrated that audience metrics are being captured in online newsrooms (e.g. MacGregor, 2007, Anderson, 2011b, Usher, 2013); many of these have also described anecdotal evidence for the importance of poor traffic statistics for the removal of individual articles, or better than expected traffic as a reason for prolonging an article's life (see e.g. Anderson 2011b, 559, MacGregor, 2007, 287). In interview, by contrast, many editors have emphasised that popularity measured in terms of readership is not a major determining factor (MacGregor, 2007, 291), and that news values remained a significant concern in terms of production and placement of news articles (Dick, 2011). Overall the impression given by journalists is one of a balance between accepting new technology whilst protecting existing news values; a response similar to those encountered by those researching more active forms of audience participation (see e.g. Thurman, 2008, Nielsen, 2013).

As Lee et al. argue, however, ethnographic research into the impact of audience

metrics on news production is methodologically problematic, for several reasons (2012, 8). Editors may misremember the impact of readership statistics on their precise decision making process, and they may also feel a need to underemphasise what is regarded by many as an essentially negative habit of placing popularity over importance in the news. Furthermore, even if ethnographic research can point to the potential importance of readership, it cannot specifically quantify this impact. For this reason, it is vital that these ethnographies are complemented by larger scale studies which measure actual journalistic practice.

Existing quantitative work on the subject is very limited. In one of the only studies of its kind, Lee et. al. (2012) explore the extent to which audience metrics, measured through the 'most read' sections of online websites, exert an effect on article positioning on the front page. Counter intuitively, they find that being part of the most read list actually had a negative impact on article positioning, which they attribute speculatively to an active resistance on the part of editors to audience influence, combined with more general dynamics of how older stories are anyway shuffled down the front page. Obviously, all news has a shelf life, and removal of old articles to make way for new ones is a basic element of the news cycle: hence the amount of time an article remains on the front page ought to have a significant positive influence on its likelihood of being removed. However within this basic limitation online editors have a degree of flexibility in deciding whether to prolong an article's life or not. Despite its usefulness, a limitation of Lee et. al.'s work is that they are not able to distinguish between the general decay of articles and how readership statistics may influence this decay, something that we seek to improve on in this study.

In addition to assessing the general impact of readership statistics, we also seek to assess whether this effect differs between different types of news. Existing research into the broader area of what is sometimes called ‘participatory journalism’ has suggested that online editors have been more willing to allow audience participation (through, for example, the submission of comments or event contribution of content) in areas of soft news such as entertainment, arts, sports, etc. (Örnebring, 2008, Jönsson & Örnebring, 2011) . Boczkowski and colleagues have published a series of research papers looking at the interactions between online audiences and journalists in the selection of news, finding a gap between the harder news choices of journalists and softer ones of readers (Boczkowski, Mitchelstein & Walter 2011; Boczkowski & Peer, 2011), subject differences between consumers’ relative clicking, emailing and commenting frequencies (Boczkowski & Mitchelstein, 2012) and that periods of heightened political activity are associated with greater focus on harder news by both journalists and consumers (Boczkowski & Mitchelstein, 2010; Boczkowski & Mitchelstein, 2012). This paper builds on this work by focusing on the visible lifetime of articles rather than their selection for publication, and by working at a larger-scale, with more continuous and more automated methods used to collect and classify the content.

Furthermore, we seek to assess the extent to which readership has differing levels of impact in different publications, in particular whether there is a ‘quality’ / ‘tabloid’ split (though this split itself is of course by no means fixed – see Deuze, 2005b). Part of the definition of tabloid style journalism lies precisely in its increased willingness to follow the demands of its audience: hence we would expect any audience effect to be stronger in tabloid style publications. More generally, as Boczkowski has argued, new

technologies take time to be integrated into the working practices of existing organisations; and the extent and manner of their integration may depend on existing organisational culture (2004). We would also expect therefore to simply observe differences across the different papers in our study.

Different outlets, different styles? Exploring the dynamics of publishing in five UK news websites

Our dataset is generated through the automatic capture of the front pages of five different major UK news outlets: the BBC, the Daily Telegraph, the Guardian, the Daily Mail and the Mirror. These sites were chosen as they are the most popular online news platforms in the UK (Press Gazette, 2013) , and also because they provide a useful balance between tabloid and quality style news outlets. We classify the BBC, the Daily Telegraph and the Guardian as ‘quality’, and the Daily Mail and the Mirror as tabloid. During the period April 18th to June 3rd we took a snapshot of the front page of each of our news outlets once every fifteen minutes, using the Heritrix web crawler (Internet Archive, 2013). This six week window provided over 20,000 front page captures, and gives us a rich and detailed picture of online publishing activity.

For this study, we recorded each individual hyperlink on each of the front pages at each time interval. After filtering out links to other websites, and links to other sections of the site (such as ‘Sports’ sub sections), we assume that every other link in our dataset is a link to an individual news article². As a result of the frequency of our sampling, we are able to say to within 15 minutes when each of these articles was first linked to from

² We also eliminate links which are only contained within the “most read” section of the site itself.

the front page, and when this link was eventually removed from the front page.

<table 1 goes here>

Table 1. Distribution and Duration of Articles

As table one shows, the average article in our dataset spends only 15 hours on the front page of any website. This average disguises some variation between sites: a typical Mirror article lasted for only 10 hours, whereas the average Daily Mail article lasts twice that amount. Variations in article duration appear to correlate closely with the total amount of articles published during our window of observation (see figure one), suggesting that the flexibility of online news has produced diverse publishing styles: some sites emphasising large front pages with (relatively) low article turnover, whilst others have smaller sites with a higher level of drop off.

<figure 1 goes here>

Figure 1. Average duration of an articles compared to total number published in the observation window. Regression line with 95% confidence intervals. Correlation

Coefficient = 0.96, $p=0.01$

In our dataset, we also draw a distinction between ‘political’ and ‘entertainment’ news, broadly following analytical categories established Bozckowski and Peer (2011, 872). These categories were based on the naming schema for web links adopted by the various different websites in question. For example, a typical article on motor racing on the BBC has a web link: <http://www.bbc.co.uk/sport/0/formula1/23085501>, specifying clearly that this is an article about sports. We classified stories about arts, culture, celebrities, fashion, travel and sports as ‘entertainment’ news (as well as a variety of other smaller subcategories). Stories about current affairs, society, health and business

were classified as more ‘political’ news. We took a restrictive stance towards our classification of political articles, hence sections which may well contain important political information (such as science and technology) were hence nevertheless classified as entertainment because we estimate that, on balance, they were more about entertainment than politics. As Delli Carpini and Williams argue (2001), this type of classification is in some ways problematic, as entertainment news can have a political quality, just as political news can be mostly about entertainment (in this context see also Baum, 2003). Nevertheless we also believe it broadly captures an important distinction.

Two outlets in our study publish appreciably more political news (the BBC and the Telegraph), whilst the Mirror is more focused on entertainment. Both the Daily Mail and the Guardian have an almost 50-50 split; perhaps slightly surprising considering the Guardian’s reputation as a ‘quality’ newspaper in comparison to the Mail’s more tabloid branding. The mean durations of these two types of news remain close to the overall mean, with the exception of the Daily Mail, whose entertainment articles remain on site significantly longer than its political news. Our initial expectation was that, as soft news is frequently perceived to be less time-sensitive, soft news pieces would likely remain longer on each news page (Boczkowski, 2009); but the Mail is the only site to lend support to this expectation.

The main independent variable in this study is the ‘popularity’ of a given article, in terms of the number of people that read it. Quantitative research on this subject has been facilitated by the widespread appearance of ‘most read’ columns on newspaper front pages, listing the most popular newspaper articles on the site at the time. As well as being of interest to researchers in their own right (Knobloch-Westerwick, Sharma, Hansen &

Alter, 2005), these areas also offer the enticing prospect of being able to measure what a newspaper's readership finds important on a given day, and have been used in a number of studies to gauge overall audience interest (see e.g. Hensing, Flaounas & Cristianini, 2012, Bozckowski & Peer. 2011, Shoemaker, Johnson, Seo & Wang, 2010, Lee et al. 2012).

Within our dataset, 12% of the articles extracted are featured at some point on the 'most read' list (as shown in table two). This number is reasonably consistent across the sites apart from at the Guardian where only 9% of articles make it on to the list and the BBC, where fully 18% of articles appear at some point. The lower position of the Guardian is likely explained by the fact that its most read list has only 5 entries, whereas most other sites have 10 (and the Daily Mail has a list of variable quantity but typically around 20).

Also of interest is the amount of time it takes for news articles to appear on the 'most read' list. While as Deszö et al. demonstrate articles typically attract most of their traffic soon after they are posted (2006), articles also need some time to attract enough traffic to make it onto the top list. In our dataset, the median time to achieving 'most read' status was two hours. This median is dragged down somewhat by the BBC, whose articles typically appear on the most read list only 15 minutes after publication (if they are to appear at all). This contrasts with the Mirror, whereby articles took an average of 4 hours to reach the most read list.

<table 2 goes here>

Table 2. Appearance of Articles in the Most Read Slot

These data are important for framing this (and future) studies which employ most

read lists. As the average article stays on the front page for 15 hours but the average time to arrival on the ‘most read’ list is only two, articles generally have a long enough life to make the ‘most read’ column while they are also live on the front page, if they are sufficiently popular. This means that most read lists do provide a reasonably accurate picture of what is currently popular on the site, rather than what was popular over the last few days.

Measuring the impact of the audience agenda

Table two also presents the mean durations of articles, categorised by whether they did or did not appear on the ‘most read’ lists at any point during our window of observation. These means provide initial support for the hypothesis that editors are influenced by the decisions of their audience, with the average ‘most read’ article lasting three hours longer than all other articles. This difference is consistent across all papers with the (surprising) exception of the Daily Mail, a point we will return to in the discussion. What the data in table two do not do, of course, is establish the direction of any effect between readership and editorial decisions: it could well be that editors are leaving well read articles up longer, or it could equally be that articles which are left up longer have more of a chance of becoming most read (though the median times to ‘most read’ status presented above suggest that the difference between 15 and 18 hours is unlikely to be decisive in achieving a high readership). In this section, we aim to test more specifically the hypothesis that readership has an influence on editorial decisions about when to remove an article from the front page.

We do this by employing a methodology drawn from event history analysis: the

Cox Proportional Hazards model (see Box-Steffensmeier & Jones, 2004). Event history analysis is a useful technique because it allows assessment of the precise characteristic we are interested in (chance of an article's 'survival') in a way that takes into account how different variables can change value during the course of measurement (in our case, how articles both appear and disappear from the most read list). The Cox model also permits us to frame our research question as follows: what impact does an article's appearance on the 'most read' list at time T have on its chance of survival to time T+1? This allows us to deal with the problems of causal ordering outlined above (whereby longer lived articles may be more likely to become "most read"). As we measure our data at fifteen minute intervals, in practice we estimate the impact on an article's chance of survival for a further fifteen minutes.

Table three presents results from five separate Cox models exploring the three elements of our research question. Each of these models takes a 'time varying' approach to the analysis, hence the dataset is separated out into 'article-period' observations, whereby each entry is a fifteen minute period in a given article's life (hence the overall N is much larger than the figures reported in tables one and two). Each entry also records whether the article was 'most read' in the previous fifteen minutes (taking into account the time lagged nature of our effect), and whether the article was removed from the front page at some point during those fifteen minutes (for further information see Mills, 2011).

Model 1 explores the overall impact of our most read variable on all the articles in the dataset, with fixed effects for 'entertainment' style news and four of our five websites (with the BBC as the reference category). This model provides further support for our hypothesis. The coefficients reported here can be interpreted in terms of a percentage

effect on the underlying hazard rate. Hence a coefficient of 0.74 for ‘most read’ articles means that the risk of such an article being removed in the next fifteen minutes is 26% lower than the risk for an article not on such a list. The confidence intervals for the effect mean that we can be 95% confident of a coefficient somewhere between 0.71 and 0.77 (in other words, of a risk lowered by between 23% and 29%). The other coefficients reported in model 1 are control variables, and reflect the differences already outlined in tables one and two.

<table 3 goes here>

Table 3. Modelling the Impact of Audience Metrics on Article Survival

Models 2 to 5 represent re-estimations of model 1 based on different subsets of our dataset. As outlined above, we expected the effect of ‘most read’ to be different for different types of news and different types of newspaper. With these models, we assess this expectation.

Models 2 and 3 focus on the difference between political and entertainment news. We hypothesised that online editors might be more willing to follow their audiences in the areas of entertainment than they would in ‘harder’ types of political news. We find a small amount of evidence for this claim, with a slightly lower coefficient for entertainment news than political news. However the confidence intervals for the two coefficients also overlap to a large degree (0.62 – 0.74 for entertainment news, and 0.67 – 0.74 for political news), meaning that it is difficult to conclude with any certainty that there is a real difference between the two categories. More importantly however, we can conclude that the effect described in model 1 clearly exists for both political and

entertainment news.

Models 4 and 5 focus on the difference between quality and tabloid papers. Here a clear difference does emerge. Contrary to expectations, the audience ‘effect’ is clearest in the quality papers. In tabloid papers, ‘most read’ status actually has a negative effect (consistent with the results reported by Lee et al., 2012), though the confidence interval comes very close to overlapping with 1 (which would be the equivalent of no effect), and the significance of this effect is much lower than the other results reported here. Hence we would be wary of arguing that tabloid editors actively reject the wishes of their audience; however we can conclude with confidence that these editors are no more likely to follow their audience than the typical ‘quality’ editor is, and in fact if anything are less likely to.

Conclusions

The political agenda has always been shaped in an important way by what the news media decide to report. In the online environment, however, news editors must decide not only what to distribute but how long it should remain prominent for. In the end, the decision about when to de-link the news may be just as significant as the decision about whether to publish it: as the amount of time an article is exposed for could be as important as the mere fact of its exposure.

One of the most important changes that online publishing has brought is large amounts of information provided to journalists and editors about the nature of their audience. With this information, professionals in the journalism business can start to construct a detailed idea of what is ‘popular’. The appearance of such statistics, combined with increasing pressures on editorial business models, has created worries about the

potential for ‘populism’ online: that editorial judgment would be overridden by traffic statistics.

This article is one of the first to systematically seek to both assess and quantify this effect. On the basis of the evidence presented here, we can be fairly confident in concluding that high traffic articles, measured through the proxy of the ‘most read’ list, do in fact spend longer in the spotlight than ones which attract less readership. On the basis of a time lagged model, and the examination of the typical time required to become most read, we can also conclude that this correlation is not solely because longer lasting articles are more likely to become most read . Furthermore, the effect appears roughly consistent between both political and entertainment style articles, meaning that this relationship is not limited just to soft types of news. Hence audience readership does have a measurable impact on the lifespan of political news.

We do however find evidence that this effect is different in different outlets. In our descriptive section, we outline diverse publishing ‘styles’, some based on a high turnover of a small amount of articles, and others a low turnover of a high amount. We also find diverse styles of editorial – readership relationship, with tabloid editors surprisingly less likely to follow their audiences than their quality counterparts. We do not have a clear explanation for this difference, though we could speculate that, as tabloid publications are already more tuned in to the wishes of their audience, the appearance of readership statistics makes less practical difference to the overall product. However it may also simply be the case that the online environment is slowly producing new journalistic practices for which the tabloid / quality distinction will be of less use.

We can conclude therefore that the audience is no longer the ignored quantity it

was in offline journalism: it has a clear impact on journalistic practice. The question which remains however is whether the extent to which editors follow their audience constitutes evidence of a new ‘populism’ in journalism; or whether it represents (as editors themselves argued above) the striking of a balance between audience needs and news values. Although this difference will always be a subjective one, our tendency is towards the latter: being ‘most read’ offers an improvement in an article’s short-run chance of survival, but it is by no means the only decisive factor, whilst non most read article still get a significant amount of ‘time in the sun’. However, while online journalism is now in its third decade of existence, it is unlikely that the practice of journalism has reached a new equilibrium. The influence of new technology will continue to filter into newsrooms; just as the technology itself continues to evolve rapidly. It is conceivable therefore that what we are witnessing here is the beginning of a new populism in journalism – though only further research will be able to establish this conclusively.

Authors

Jonathan Bright is a Research Fellow at the Oxford Internet Institute.

jonathan.bright@oii.ox.ac.uk

Tom Nicholls is a doctoral candidate at the Oxford Internet Institute.

tom.nicholls@oii.ox.ac.uk

Mailing address for both authors:

1 St Giles' Oxford, United Kingdom, OX1 3JS

References

- Anderson, C. W. (2011a). Deliberative, Agonistic, and Algorithmic Audiences: Journalism's Vision of its Public in an Age of Audience Transparency. *International Journal of Communication*, 5, 529-547.
- Anderson, C. W. (2011b). Between creative and quantified audiences: Web metrics and changing patterns of newswork in local US newsrooms. *Journalism*, 12(5), 550-566.
- Baum, M. A. (2003). Soft news and political knowledge: evidence of absence or absence of evidence?. *Political Communication*, 20, 173-190.
- Bozckowski, P. J. (2004). *Digitizing the News*. Cambridge: The MIT Press.
- Bozckowski, P. J. (2009). Rethinking Hard and Soft News Production: From Common Ground to Divergent Paths. *Journal of Communication*, 59(1), 98-116.
- Boczowski, P. J., & Mitchelstein, E. (2010). Is There a Gap Between the News Choices of Journalists and Consumers? A relational and dynamic approach. *The International Journal of Press/Politics*, 15(4), 420-440.
- Boczowski, P. J., & Mitchelstein, E. (2012). How Users Take Advantage of Different Forms of Interactivity on Online News Sites: Clicking, E-Mailing, and Commenting. *Human Communication Research*, 38(1), 1-22.
- Boczowski, P. J., Mitchelstein, E., & Walter, M. (2011). Convergence across divergence: Understanding the gap in the online news choices of journalists and consumers in Western Europe and Latin America. *Communication Research*, 38(3), 376-396.
- Bozckowski, P. J., & Peer, L. (2011). The Choice Gap: The Divergent Online News

Preferences of Journalists and Consumers. *Journal of Communication*, 61(5), 857-876.

Box-Steffensmeier, J. M., & Jones, B. S. (2004). *Event History Modeling: A Guide for Social Scientists*. Cambridge: Cambridge University Press.

Brossard, D., & Scheufele, D. A. (2013). Science, New Media and the Public. *Science*, 339, 40-41.

Bruns, A. (2008). The Active Audience : transforming journalism from gatekeeping to gatewatching. In C. Paterson & D. Domingo (Eds.), *Making Online News*. Peter Lang: New York.

Chung, D. S. (2007). Profits and Perils: Online News Producers' Perceptions of Interactivity and Uses of Interactive Features. *Convergence*, 13(1), 43-61.

Cohen, E. L. (2002). Online journalism as market driven journalism. *Journal of Broadcasting and Electronic Media*, 46(4), 532-548.

Delli Carpini, M. X., & Williams, B. A. (2001). Let Us Infotain You: Politics in the New Media Age. In: W. L. Bennett & R. M. Entman (Eds.), *Mediated Politics: Communication in the Future of Democracy*. Cambridge: Cambridge University Press.

Deuze, M. (2003). The Web and its Journalisms: Considering the Consequences of Different Types of Newsmedia Online. *New Media and Society*, 5(2), 203-230.

Deuze, M. (2005a). What is journalism? Professional identity and ideology of journalists reconsidered. *Journalism*, 6, 442-464.

Deuze, M. (2005b). Popular journalism and professional ideology: tabloid reporters and

- editors speak out. *Media, Culture and Society*, 27(6), 861-882.
- Deuze, M. (2008). The changing context of news work: Liquid journalism and monitorial citizenship. *International Journal of Communication*, 2(5), 848-865.
- Deszö, Z., Almaas, E., Lukács, A., Rácz, B., Szakadát, I., & Barabási, A.-L. (2006). Dynamics of Information Access on the Web. *Physical Review E*, 73(6).
- Dick, M. (2011). Search Engine Optimisation in UK News Production. *Journalism Practice*, 5(4), 462-477.
- Domingo, D. (2008). Interactivity in the daily routines of online newsrooms: Dealing with an uncomfortable myth. *Journal of Computer-Mediated Communication*, 13(3), 680-704.
- Harcup, T., & O'Neill, D. (2001). What is News? *Journalism Studies*, 2(1), 261-280.
- Hensinger, E., Flaounas, I., & Cristianini, N. (2012). Modelling and predicting news popularity. *Pattern Analysis and Applications*.
- Internet Archive (2013). *Heritrix 3.1.1*. Computer software. Available from:
<http://builds.archive.org:8080/maven2/org/archive/heritrix/heritrix/3.1.1/>
- Jönsson, A.M., & Örnebring, H. (2011). User-generated content and the news: Empowerment of citizens or interactive illusion? *Journalism Practice*, 5(2), 127-144.
- Karlsson, M., & Strömbäck, J. (2010). 'Freezing the Flow of Online News. Exploring Approaches to the Study of the Liquidity of Online News', in: *Journalism Studies*, 11(1), 2-19.
- Knobloch-Westerwick, S., Sharma, N., Hansen, D. L., & Alter, S. (2005). Impact of Popularity Indications on Readers' Selective Exposure to Online News.

Journal of Broadcasting and Electronic Media, 49(3), 296-313.

Lee, A. M., Lewis, S. C., & Powers, M. (2012). Audience Clicks and News Placement: A Study of Time-Lagged Influence in Online Journalism. *Communication Research*, XX(X), 1-26.

Loosen, W., & Schmidt, J-H. (2012). (Re-)discovering the audience. *Information, Communication and Society*, 15(6), 867-887.

MacGregor, P. (2007). Tracking the Online Audience. *Journalism Studies*, 8(2), 280-298.

McCombs, M. (2004). *Setting the Agenda : The Mass Media and Public Opinion*. London: Polity Press.

Mills, M. (2011). *Introducing Survival and Event History Analysis*. London: SAGE.

Napoli, P. M. (2010). *Audience Evolution: New Technologies and the transformation of the media audience*. New York: Columbia University Press.

Newman, N., Dutton, W. H., & Blank, G. (2012). Social Media in the Changing Ecology of News: The Fourth and Fifth Estates in Britain. *International Journal of Internet Science*, 7(1), 6-22.

Nielsen, C. E. (2013). Coproduction or cohabitation: Are anonymous online comments on newspaper websites shaping news content. *New Media and Society*, forthcoming

Oblak, T. (2005). The lack of interactivity and hypertextuality in online media. *Gazette*, 67(1), 87-106.

Örnebring, H. (2008). The consumer as producer – of what? *Journalism Studies*, 9(5), 771-785.

Paterson, C., & Domingo, D. (Eds.) . (2008). *Making Online News*. Peter Lang: New

York.

Paulussen, S., Ugille, P. (2008). User Generated Content in the Newsroom: Professional and Organization Constraints on Participatory Journalism. *Westminster Papers in Communication and Culture*, 5(2), 24-41.

Ponsford, D. (2013, May 23). Telegraph website traffic continues to grow despite launch of metered paywall. *Press Gazette*. Retrieved June 28, 2013, from:
<http://www.pressgazette.co.uk/adoption-metered-paywall-has-little-effect-telegraphs-overall-website-traffic>

Shoemaker, P. J., & Vos, T. P. (2009). *Gatekeeping Theory*. Abingdon: Routledge.

Shoemaker, P. J., Johnson, P. R., Seo, H., & Wang, X. (2010). Readers as gatekeepers in online news: Brazil, China and the United States. *Brazilian Journalism Research*, 6(1), 55-77.

Thurman, N. (2008). Forums for citizen journalists? Adoption of user generated content initiatives by online news media. *New Media and Society*, 10(1), 139-157.

Usher, N. (2013). Al-Jazeera English Online: Understanding Web metrics and news production when a quantified audience is not a commodified audience. *Digital Journalism*, forthcoming

Tables and Figures

Paper	Articles	Political News	Entertainment News	Mean Duration (SD)	Mean Political Duration	Mean Entertainment Duration
BBC	6745	76%	24%	12 (10)	12	12
Daily Mail	14484	52%	48%	20 (13)	16	24
Guardian	9380	55%	45%	13 (11)	14	13
Mirror	5075	40%	60%	10 (10)	11	9
Telegraph	7296	60%	40%	14 (11)	14	15
Total	42980	56%	44%	15 (12)	14	17

Table 1. Distribution and Duration of Articles

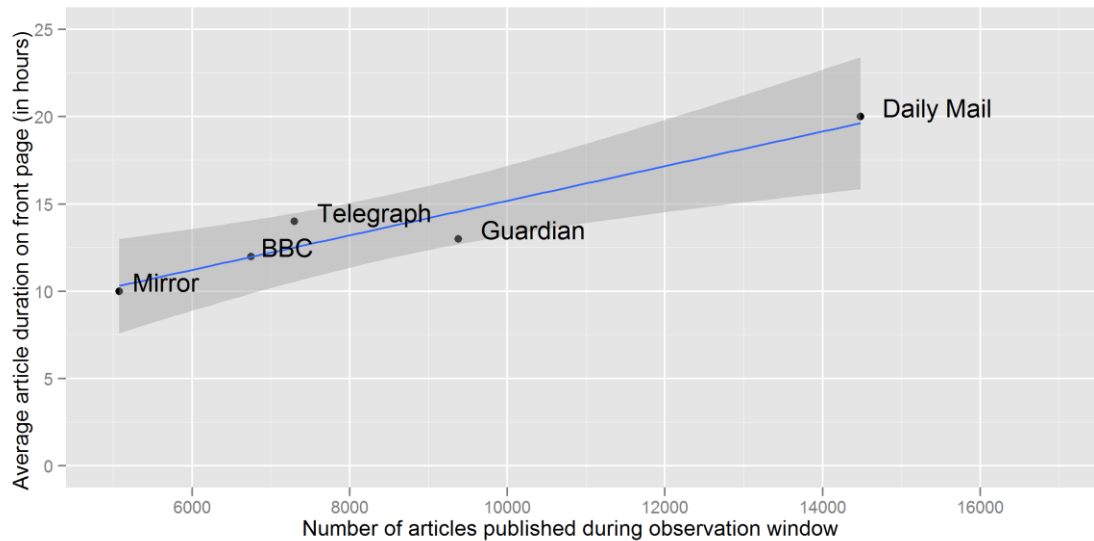


Figure 1. Average duration of an articles compared to total number published in the observation window. Regression line with 95% confidence intervals. Correlation

Coefficient = 0.96, $p=0.01$

Paper	Articles	Most	Most Read	% Most	Median	Mean	Mean
-------	----------	------	-----------	--------	--------	------	------

		Read Slots		Read	time to Most Read (hours)	Duration (most read articles)	Duration (all other articles)
BBC	6745	10	1222	18%	0.25	18	11
Daily Mail	14484	Up to 20	1593	11%	2.3	17	20
Guardian	9380	5	805	9%	2.8	18	13
Mirror	5075	10	593	12%	4	14	9
Telegraph	7296	10	711	10%	2.5	20	14
Total	42980		4924	12%	2	18	15

Table 2. Appearance of Articles in the Most Read Slot

	Model 1 – Overall impact of Most Read	Model 2 – Political News	Model 3 – Entertainment News	Model 4 – Quality Papers	Model 5 – Tabloid Papers
Most Read at time T	0.74 (0.02) ***	0.70 (0.03) ***	0.68 (0.05) ***	0.54 (0.03) ***	1.10 (0.03) **
Entertainment	0.89 (0.01) ***			0.91 (0.01) ***	0.79 (0.01) ***
Daily Mail	0.49 (0.01) ***	0.58 (0.02) ***	0.41 (0.03) ***		
Guardian	0.64 (0.02) ***	0.61 (0.02) ***	0.66 (0.03) ***		
Mirror	1.32 (0.02) ***	1.06 (0.02) *	1.57 (0.03) ***		
Telegraph	0.70 (0.02) ***	0.68 (0.02) ***	0.72 (0.03) ***		
N	2,429,940	1,247,210	1,182,730	1,147,878	1,282,062

Table 3. Modelling the Impact of Audience Metrics on Article Survival