



News selection criteria in the digital age: Professional norms versus online audience metrics

Journalism

1–17

© The Author(s) 2015

Reprints and permissions:

sagepub.co.uk/journalsPermissions.nav

DOI: 10.1177/1464884915595474

jou.sagepub.com



Kasper Welbers

VU University Amsterdam, The Netherlands

Wouter van Atteveldt

VU University Amsterdam, The Netherlands

Jan Kleinnijenhuis

VU University Amsterdam, The Netherlands

Nel Ruigrok

Erasmus University Rotterdam, The Netherlands

Joep Schaper

University of Amsterdam, The Netherlands

Abstract

On newspaper websites, journalists can observe the preferences of the audience in unprecedented detail and for low costs, based on the audience clicks (i.e. page views) for specific news articles. This article addresses whether journalists use this information to cater to audience preferences in their news selection choices. We analyzed the print and online editions of five national newspapers from the Netherlands with a mixed-method approach. Using a cross-lagged analysis covering 6 months, we found that storylines of the most-viewed articles were more likely to receive attention in subsequent reporting, which indicates that audience clicks affect news selection. However, based on interviews with editors we found that they consider the use of

Corresponding author:

Kasper Welbers, Faculty of Social Sciences, VU University Amsterdam, De Boelelaan 1081, 1081 HV Amsterdam, The Netherlands.

Email: k.welbers@vu.nl

this information for news selection to conflict with professional norms. We elaborate on the implications of this discrepancy in the norms and behaviors of journalists, and project directions for future studies.

Keywords

Gatekeeping, journalism, online audience metrics, professional norms

Introduction

Journalists fulfill an important role in society as gatekeepers: they filter and transform the set of daily world events into a limited number of news articles (Shoemaker, 1991: 1). This filtering process has a strong influence on the social realities of citizens (Shoemaker and Vos, 2009). Accordingly, the criteria used in this filtering process, and whether these criteria are beneficial or detrimental to society, are issues that have concerned scholars for many decades (Cook, 2005; Schudson, 2003).

A particularly relevant point of discussion is to what extent the preferences of the audience should be taken into the equation. A central element in this debate is a continuum of two opposing views on the role of journalism in society. Put briefly, should journalists determine what news is based on their professional judgment – a *trustee model* – or should they let the audience decide what it wants – a *market model* (Schudson, 2003)? Recently, this discussion has been revived due to the increasing popularity of the Internet as a news medium. Online, journalists can track how many times a web page is viewed (from here on referred to as audience clicks), which discloses real-time information about the interests of the audience in specific news articles. In this article, we study whether and how newspapers in the Netherlands monitor audience clicks on their websites, and whether this affects their news selection choices.

Several studies have found effects of online audience metrics such as audience clicks on journalistic decision-making (Anderson, 2011; Lee et al., 2014; Vu, 2014), but there is still much left to explore and the situation is currently still developing. Lee et al. (2014) specifically called for more quantitative studies, arguing that ‘The challenge for researchers [...] is to move beyond self-reports of journalistic perception and behavior, and instead use quantitative methods that reveal a more precise, longitudinal rendering of the relationship between audience behaviors and editorial decisions’ (p. 17).

We contribute to the literature in three ways. First, we build on the quantitative approach used by Lee et al. (2014), using the most-viewed articles lists on newspaper websites, to analyze the influence of audience clicks for five newspaper organizations. Where Lee et al. analyzed the effect of audience clicks on news placement, we focus on news selection: does news with many clicks receive more future coverage? There are studies that suggest that this is the case (Anderson, 2011; Vu, 2014), but this has not yet been confirmed based on this type of quantitative approach. Also, most literature focuses on the context of the United States. We focus on the context of the Netherlands, which has a different media system (Hallin and Mancini, 2004).

Second, we conducted interviews with editors of the newspapers in our study to inquire how they monitor and use audience clicks. This enables us to compare our quantitative

findings to self-reports of journalists. Lee et al. (2014) noticed a discrepancy between how journalists say that they use audience clicks and how they are actually influenced by them. This discrepancy can be attributed to the struggle journalists engage in to balance the use of online audience metrics with their professional norms (Anderson, 2011). Based on our mixed-method approach, we analyze this discrepancy and discuss this struggle.

Finally, we analyze the influence of audience clicks on both the online and print editions of newspapers. Audience clicks enable online editors to cater to the preferences of their audience, but for print editors this information also provides more insight into audience preferences than previously possible. Being part of the same organization, this information could easily be exchanged. However, managing this exchange would be an explicit acknowledgment of the importance of audience preferences in news selection. Also, print journalists might have a more traditional 'vision of the audience' compared to online journalists (Anderson, 2011) and therefore pose more normative resistance. Analyzing the influence of audience clicks on print news can therefore be considered a more rigorous test of the lure of online audience metrics.

Role of audience preferences in gatekeeping

The gatekeeping metaphor was first used in the context of news selection by White (1950), who observed how a single newspaper editor decided which stories were published. The question addressed by White (1950) was, 'how does the gatekeeper make this selection, what are the factors involved in this process?' Subsequent studies further explored these factors, looking not only at individual preferences of journalists but also taking technological, cultural, organizational, and institutional factors into account (Shoemaker and Vos, 2009).

Monitoring online audience metrics can affect the gatekeeping process because it enables journalists to more accurately take the *preferences* of the audience into account. However, the role of audience preferences in gatekeeping is debatable because journalists also have a responsibility to select the information that is *important* for the audience to know. Thus, journalists have to balance news that they think is important to the audience with news that the audience prefers to consume (Gans, 1979).

How journalists think this should be balanced relates to their normative conception of the role of journalism in society. Schudson (2003) discusses two relevant conflicting normative frameworks: the Trustee Model and the Market Model. In the Trustee Model, journalists prioritize their professional judgment to decide what news is important for citizens to know. In contrast, in the Market Model, journalists prioritize catering to the preferences of the audience (i.e. the demands of the market).

The Trustee Model dovetails with the ideals of representative democracy, which dictates that the public should be given the means to educate themselves in order to acquire an 'enlightened understanding' of political and societal matters (Dahl, 1998). The main critique on the model is that by disregarding the preferences of their audiences journalists have become too disconnected from public life and too caught up in elitist values (Deuze, 2008). Notwithstanding this critique of the model, it appears that, as Paulussen et al. (2007) observe, 'it still dominates professional and scholarly literature on journalism, and defines the conventional framework of journalism education' (p. 134).

Despite this dominance, since the 1980s scholars have observed ‘a move away from reliance on craft norms defining what is newsworthy and how to report, toward a journalism based on serving the marketplace’ (McManus, 1995: 301s). The rise of this Market Model is intertwined with the increased competition in the news market (Hamilton, 2004). Normatively, the Market Model addresses the critique on the Trustee Model by giving the audience what it wants, instead of what elitist journalists think they ought to want (Fancher, 1987). In return, the main critique on the Market Model is that it might give the audience what it wants, but this is not necessarily what it needs (Anderson and Ward, 2007).

There are differences in the extent to which newspaper organizations have shifted toward a Market Model, in particular across the divide of ‘popular’ and elite, or ‘quality’, newspapers (Bakker and Scholten, 2013). Popular newspapers focus more on soft news (e.g. entertainment, lifestyle), whereas quality newspapers focus more on hard news (public affairs, particularly involving the government; Ryfe, 2012). Soft news is generally more popular than hard news (Singer, 2011). The focus of popular newspapers on soft news is more compatible with the Market Model.

The lure of online audience metrics

With real-time information on the page views of specific news stories, journalists can observe the preferences of the audience in unprecedented detail. However, since the use of audience preferences conflicts with the norms of the Trustee Model, the question is whether journalists will also actually use this information. In the past, there were also various ways to measure audience preferences, and since the 1980s it has been common for newspapers to enlist consultants to conduct market research (Hamilton, 2004; Ryfe, 2012), but the influence of audience preferences nevertheless remained limited due to this normative resistance (Deuze, 2008; Singer, 2011). As Anderson (2011) concluded, ‘Journalism’s professional self-conception and its somewhat paternalistic vision of its audience have historically served to blunt the impact of whatever limited audience measurement technologies might influence news production’ (p. 554).

There are two main factors that, when combined, might override the normative resistance against market research: the advantages of the new technology and the increasing economic pressures of the news market. Gans (1979) predicted that ‘At present, executive and professional intuition remains dominant; but should commercial considerations become more urgent, the [market] researchers may find themselves with sufficient funds to demonstrate whether they can enhance the corporate balance sheet’ (p. 234). As it turns out, not only *did* commercial considerations become more urgent (Hamilton, 2004; McManus, 1995; Ryfe, 2012), but research has also become much more effective and less expensive. In the past, relatively expensive research could only report on ‘recurring patterns’ in news consumption behavior, which was less useful to journalists who must decide on ‘individual stories’ (Gans, 1979: 232). At present, interests in individual stories can be monitored directly, in real-time, and for moderate costs. Through the combination of these economical and technological factors, the integration of audience interests in the gatekeeping process has become much more alluring.

Several studies have indeed found that online audience metrics, and specifically audience clicks, affect journalistic decision-making. Based on an ethnographic newsroom study, Anderson (2011) concluded that the use of audience metrics affects the process of ‘deciding what’s news’ (p. 563). Lee et al. (2014) found, by means of a cross-lagged analysis, that audience clicks affected the placement of articles on the web page. Vu (2014) conducted a survey among American newspapers and news sites and found that ‘to some extent, editors are willing to adjust their editorial decision-making based on web metrics’ (p. 11). Based on these results, we expect audience clicks to affect the news selection choices of journalists in the Netherlands:

H1. Audience clicks influence the news selection process in the Netherlands.

We specify the hypothesis to the context of the Netherlands because the influence of audience clicks is likely to differ depending on a country’s media environment, in particular regarding competition and professionalism (see, for example, Usher, 2013). Most previous research on audience clicks focuses on the United States (e.g. Anderson, 2011; Lee et al., 2014; Vu, 2014) where competition and professionalism are high (Hallin and Mancini, 2004). Compared to the United States, the media system of the Netherlands is also highly professionalized, but is less market dominated (Brants and Van Praag, 2006; Hallin and Mancini, 2004). Importantly, however, market pressure is likely to have increased since newspaper organizations suffered greatly from the financial crisis and recession in the past decade (Bakker and Scholten, 2013). Overall, we thus suspect a high level of competition to attract audiences, for which monitoring audience clicks can be an effective – and therefore luring – practice. But also, we expect strong normative resistance due to a high level of professionalism (Anderson, 2011).

Journalism as an institution

The rise of online audience metrics puts journalists in a tough position. On the one hand, they are pushed to develop routines to use these metrics, mainly due to the increasing influence of market logic (Gans, 1979; Hamilton, 2004; Vu, 2014). On the other hand, they are pulled back by traditional norms of professional journalism (Anderson, 2011). Consequently, journalists are struggling to balance the use of online audience metrics with traditional norms (Anderson, 2011; Lee et al., 2014; Usher, 2013).

In order to understand this struggle, it is important to study it not only at the level of individual journalists and newsrooms but also at the level of journalism as an *institution*. Journalists across news organizations are highly similar and interdependent as members of the same profession (Cook, 2005). Accordingly, the norms for whether and how online audience metrics *should* be used by professional journalists cannot be resolved by individual news organizations.

A defining quality of institutions is that they have social patterns that are taken for granted (i.e. ‘valued in and of themselves’) as the ways in which certain actors *should* behave (Cook, 2005: 67). The norms of professional journalism are such social patterns, and they are valuable to journalism because they define what it means to be a professional journalist. Journalism is not a typical profession, such as medicine or law; there

are no formal rules and qualifications that one has to follow and acquire in order to be a journalist. For journalists, this is troublesome because it leaves them uncertain about the quality of their work (what is news?), and it threatens their professional self-conception (what is a journalist?). The reason that journalism can nevertheless be considered to be a profession is that over time journalists have developed certain norms for what news is and how news should be made. These norms have become taken for granted across news organizations as the ways in which news *should* be made, and thereby infused with power similar to formal rules and qualifications (Cook, 2005; Gans, 1979).

A consequence of these shared norms is that journalism is slow to adapt to changes. Institutions have a 'stickiness' quality: they 'evolve in a path-dependent pattern. Because initial patterns have positive-feedback characteristics, actors have an incentive to adapt to prevailing conditions rather than seek to change the institutional order' (Ryfe, 2006: 137). This stickiness of institutions can serve as a theoretical explanation for *why* traditional journalistic norms have blunted the influence of audience research in the past (Anderson, 2011) and more generally the resistance against the Market Model. This invites us to think about this resistance as not only based on individual beliefs, but as an institutional force with supra-individual purpose and consequences.

Social patterns of actors within institutions tend to evolve in tandem and to converge, and the norms of professional journalism have indeed evolved as such (Cook, 2005; Schudson, 2003). At present, we still know little about the similarities and differences between news organizations in how they think about the use of online audience metrics, how they use them, and how this affects the news production process. We therefore address the following question:

RQ1. How do different news organizations monitor and use audience clicks, and does this relate to the influence of audience clicks on their news selection choices?

Method

We conducted a content analysis and interviews, which we will discuss in order.

Content analysis: Data

We analyzed the print and online editions of five national newspapers from the Netherlands from 12 January 2013 to 30 July 2013. A representative sample of the market research agency GfK in July 2012 (n=1805) shows that over the course of 1 week, 46.4 percent attended to the print edition of at least one of these five newspapers and 31.9 percent attended to at least one of the online editions.

Two newspapers in our sample, *De Telegraaf* and *Algemeen Dagblad*, are 'popular' newspapers, and the other three, *de Volkskrant*, *Trouw*, and *NRC Handelsblad*, are elite or 'quality' newspapers (Bakker and Scholten, 2013). All data were collected using web scraping algorithms.¹

The viewing behavior of the audience was measured using the top 5 most-viewed articles, which is available on the website of each newspaper in our sample. A computer script was used to gather the top 5 articles twice per day.² Due to occasional failure of the script, the top 5 articles for some days were not collected correctly. Since failure of the

Table 1. Average daily number of articles in our data per medium.

| Newspaper | Type | Average articles per day | | | |
|-------------------------|---------|--------------------------|-------|--------|-------|
| | | Days | Print | Online | Top 5 |
| <i>Algemeen Dagblad</i> | Popular | 76 | 139.4 | 238.3 | 4.4 |
| <i>De Telegraaf</i> | Popular | 124 | 167.8 | 215.3 | 5.2 |
| <i>de Volkskrant</i> | Quality | 117 | 118.1 | 188.1 | 5.8 |
| <i>NRC Handelsblad</i> | Quality | 128 | 116.6 | 31.4 | 5.5 |
| <i>Trouw</i> | Quality | 122 | 92.7 | 136.5 | 5.1 |

script is unlikely to be related to the event addressed in an article, this introduces noise in the data, but is unlikely to bias the results. Days on which there are less than three top 5 articles are not included as cases in the analysis. Also, data for *Algemeen Dagblad* was not collected for the first 3 months (implications of this missing data are addressed in the ‘Results’ section). Table 1 summarizes the data for each newspaper.

Content analysis: Operationalization

News articles were coded into categories so that it can be analyzed whether audience clicks for a specific news article cause more attention for news of the same category. We used *storylines* as categories, defined as all news articles addressing the same or related events. The relation between events has to be concrete (e.g. specific people, locations, developments) and not only abstract (e.g. type of event). For instance, we use the Trayvon Martin murder case and the war in Syria, and not murder cases or war in general. This more specific level of coding is appropriate given the level of specificity at which audience clicks inform journalists about audience preferences.

To measure the attention for a storyline, a search query was created based on its most distinguishing words and word combinations. This search query was then used to code which articles address the storyline. If a storyline has a clear starting or ending date, we only analyzed it within this period.

The storylines and search queries were manually selected and formulated, but to explore what storylines were prominent and what words to use to formulate the search queries, we first performed an exploratory text analysis using Latent Dirichlet allocation (LDA). LDA automatically classifies articles based on word clusters that represent topics (Blei et al., 2003). We performed LDA to extract 50 topics from the top 5 articles. We only used a selection of the words in these documents as variables: only nouns, names, and verbs that occur in at least three articles and that occurred in less than 50 percent of the articles. Based on the 50 topics, we manually selected 21 storylines and used the word clusters as inspiration for search queries.³

Content analysis: Model specification

To analyze the influence of audience clicks on news selection, we conducted cross-lagged regression analyses. This is a common method to analyze causal relations of time-series (i.e. measurements over time). The philosophy of this method is to investigate a causal

relation by measuring two key elements of predictive causality: the cause needs to occur before the effect and the cause needs to predict the effect controlled for other potential causes (Granger, 1988). Using a regression model, we estimate to what extent lagged (i.e. previous) measurements of audience clicks for a certain storyline predict newspaper coverage of this storyline.

We control for auto-regression (i.e. the lagged dependent variable) because a measurement in a time-series often depends on previous measurements (Hollanders and Vliegthart, 2008). For news coverage specifically, it is widely observed that recent attention for an event increases the likelihood that related events are covered (Galtung and Ruge, 1965). In addition, we control for cross-lagged effects between print and online news. Print and online editions of newspapers are likely to influence each others' content (Boczkowski and De Santos, 2007) and can both influence audience clicks. By controlling for cross-lagged effects between print and online news, we account for spurious effects.

We used separate models for print and online newspapers and audience clicks. The dependent variable in each model is the number of articles about the storyline per day. Since this is a count variable, we used generalized linear models with a Poisson link function (see, for example, Vliegthart and Walgrave, 2008). We pooled different storylines together. To account for differences in the average number of articles per storyline, we used multilevel models in which observations (days) are nested in storylines. The independent variables are the lagged scores for the print edition, online edition, and audience clicks. To analyze different effects between newspapers, we included dummy variables for each newspaper and interaction effects for each dummy variable with all three independent variables.

There are some complications to determine how to calculate the lagged scores. First, the print news is published at a fixed time in morning,⁴ whereas the online news and audience clicks occur throughout the day. Also, because we use days as discrete time units, and thereby ignore the temporal order of online news and audience clicks, their auto-regression and cross-lagged effects can only be calculated across days. We therefore only measure whether online news and audience clicks are influenced by the online news and audience clicks of the previous day and by the print news of that morning. We do think that most effects occur after one or several days because future coverage of a specific storyline also depends on whether new information is available. Still, effects can occur within the same day, so it should be taken into account that we might underestimate the influence of audience clicks.

Second, there is virtually no theoretical understanding for the time span within which audience clicks and news selection affect each other. It makes sense that influence lasts longer than 1 day but decays over time.⁵ We therefore used a moving window approach with exponential decay: the lag is calculated as the sum of the past 7 days, but days further back in time are weighted down, with a decay constant of 0.5 (half-life = 1.39 days).⁶

Interviews

We interviewed an editor of the online edition of each newspaper. We focus on the online editions because we assume that online editors are most knowledgeable about whether

and how online traffic is monitored. For the three quality newspapers, we also interviewed an editor of the print editions. Since we could not interview print editors of the popular newspapers, we only use these interviews for additional interpretation.

All interviews were conducted by the same interviewer in the period between 16 September 2013 and 3 October 2013. The setting was face-to-face, except for *De Telegraaf* and *Algemeen Dagblad*, which only agreed to interviews by telephone. The style was conversational (i.e. adapting to the story of the interviewee), but two questions were central for each news organization: Whether they monitor audience clicks and whether this information is used to focus news selection on popular news.

Results

First, we present the findings from our cross-lagged analysis. Second, we present the statements of editors regarding the monitoring and use of audience clicks.

Influence of audience clicks on news selection

Table 2 presents the results of the cross-lagged analysis.⁷ We first look at model 1, which shows the effects on news selection in print news. The effect of lagged print news (i.e. auto-regression) and the effect of lagged online news serve as control variables. What we are interested in is whether editors are more likely to publish articles about storylines that were recently in the top 5 most-viewed articles, which is measured with the effect of *audience clicks*. This effect is positive and significant for *Algemeen Dagblad* ($b=0.845$, $p<0.001$), *De Telegraaf* ($b=0.166$, $p<0.05$), *de Volkskrant* ($b=0.308$, $p<0.001$), and *NRC Handelsblad* ($b=0.444$, $p<0.001$). Only the print edition of *Trouw* was not influenced by audience clicks ($b=0.162$, $p=ns$).

Second, we look at model 2 for the website. Here, the effect of audience clicks is again positive and significant for *Algemeen Dagblad* ($b=0.505$, $p<0.001$), *De Telegraaf* ($b=0.133$, $p<0.05$), and *de Volkskrant* ($b=0.123$, $p<0.01$). Audience clicks had no effect on the website of *NRC Handelsblad* ($b=0.121$, $p=ns$) and again had no effect on *Trouw* ($b=-0.041$, $p=ns$). Note that since this cross-lagged model does not take influence within the same day into account, we might underestimate the effects.

We analyzed the effect of audience clicks on print editions of newspapers because we think this is a more rigorous test compared to online editions: giving print editors access to audience clicks requires clear top-down management and might receive more normative resistance. It is surprising, therefore, that there was a significant effect of audience clicks on the print news of *NRC Handelsblad* but not on the online news. What might explain this is that for print editors, information about audience clicks has an additional advantage. For online newspapers, audience clicks only inform editors about audience interests in news they already covered. For print newspapers, audience clicks can in addition be used to determine which of the articles covered online during the day should be covered in tomorrow's newspaper.

Audience clicks for storylines were influenced by either or both print and online news, which indicates that it is important to control for cross-lagged effects between print and online news in models 1 and 2 to account for spurious effects. Also, for all

Table 2. Cross-lagged regressions among the print edition, website, and audience clicks for each medium.

| Variables | Model 1 | Model 2 | Model 3 |
|--|------------|-------------|-----------------|
| | Print news | Online news | Audience clicks |
| Constants | | | |
| <i>Algemeen Dagblad</i> | -0.787*** | 0.279** | -3.343*** |
| <i>De Telegraaf</i> | -0.811*** | -0.383*** | -2.899*** |
| <i>de Volkskrant</i> | -0.763*** | -0.268** | -2.711*** |
| <i>NRC Handelsblad</i> | -0.953*** | -1.51*** | -2.795*** |
| <i>Trouw</i> | -0.999*** | -0.457*** | -3.09*** |
| Effects | | | |
| <i>Algemeen Dagblad</i> | | | |
| Print | 0.019 | 0.037*** | 0.03 |
| Online | 0.042*** | 0.044*** | 0.035* |
| Audience clicks | 0.845*** | 0.505*** | 1.51*** |
| <i>De Telegraaf</i> | | | |
| Print | 0.099*** | 0.049*** | 0.053* |
| Online | 0.068*** | 0.124*** | 0.067** |
| Audience clicks | 0.166* | 0.133* | 0.814*** |
| <i>de Volkskrant</i> | | | |
| Print | 0.09*** | 0.157*** | 0.188*** |
| Online | 0.041*** | 0.043*** | 0.018 |
| Audience clicks | 0.308*** | 0.123** | 0.357** |
| <i>NRC Handelsblad</i> | | | |
| Print | 0.179*** | 0.072*** | 0.088* |
| Online | 0.014 | 0.29*** | 0.157 |
| Audience clicks | 0.444*** | 0.121 | 0.603** |
| <i>Trouw</i> | | | |
| Print | 0.085*** | 0.163*** | 0.265*** |
| Online | 0.087*** | 0.076*** | 0.019 |
| Audience clicks | 0.162 | -0.041 | 0.109 |
| BIC | 18,755 | 24,588 | 4765 |
| Log likelihood | -9283 | -12,199 | -2288 |
| LLbase-model $\Delta\chi^2$ ^a | 501.98*** | 571.78*** | 136.44*** |

BIC: Bayesian information criterion.

^aModel improvement compared to base with only intercepts and auto-regression.

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

newspapers except *Trouw*, there were auto-regression effects. This indicates that audiences are more likely to click on a storylines if it recently received many clicks, which might be explained by the increased prominence of the storyline (Lee et al., 2014).

Note that since the first 3 months of *Algemeen Dagblad* are missing, it has to be taken into account that the results are not directly comparable to the other newspapers. Also,

Table 3. Statements by website editors regarding the monitoring and use of audience clicks.

| | Are audience clicks monitored? | Do audience clicks affect news selection? |
|-------------------------|---|--|
| <i>De Telegraaf</i> | 'Throughout the entire day' | 'At most it affects where on the page the article is placed' |
| <i>Algemeen Dagblad</i> | 'Always. We are constantly aware of what visitors click on' | 'If important news does not score we do bring it, but on the [less prominent] side' |
| <i>de Volkskrant</i> | 'That is an integral part of the news coordination of a site' | The first consideration is what people have to know, after that you can 'keep track of what is clicked' and 'turn knobs' |
| <i>NRC Handelsblad</i> | 'All editors have access to the data of visitor-numbers' | For news placement we do look at the number of page views, 'but it does not play a role in news selection' |
| <i>Trouw</i> | 'I certainly do, and by now I know what is clicked on' | 'What counts for us are clicks on articles that fit <i>Trouw</i> ' |

this might affect how the model converges. We therefore also performed the analysis without *Algemeen Dagblad* to confirm that this does not relevantly affect the findings for the other newspapers.⁸

Overall, these results support our hypothesis that audience clicks affect news selection. A minority of the news organizations resisted the use of audience clicks. In the next section, we discuss the interviews and elaborate on this resistance.

Statements from editors

In Table 3, we summarize the main statements of the editors of the websites about the monitoring and use of audience clicks. Two common observations are especially evident. The first is that each news organization actively monitors audience clicks. The second common observation is that all news organizations deny or strongly nuance the influence of audience clicks on news selection. The main use of audience clicks is argued to be for news placement or for testing and revising headlines. This contradicts some of our quantitative findings, and assuming our analysis to be correct implies that journalists are either unaware of the influence of audience clicks on news selection or unwilling to admit it. Lee et al. (2014) similarly found that editor denials of the influence of audience clicks on news placement did not match their findings. This contradiction between what journalists say and what they do is relevant to our understanding of the struggle journalists are engaged in to balance the use of audience clicks with traditional norms. For a more detailed insight into this struggle, we elaborate on some of the differences between news organizations.

We first discuss the differences in how news organizations manage the monitoring of audience clicks. The most coordinated pursuit of audience clicks was observed at *De Telegraaf*, where journalists had access to screens in the newsroom that keep track of audience clicks. The editors of the print edition and website also share the same

newsroom, so the visibility of audience clicks is managed top-down throughout the organization.

In the other news organizations, the print and online newsrooms do not share the same space, so sharing insights from audience clicks requires more active involvement of the editors. Also, there was much less top-down communication of this information. For *Trouw*, this consisted of ‘mailing a weekly report. [...] Purely figures, no advice’. The editor of the website of *de Volkskrant* also mentioned the use of email to communicate audience metrics. The editor of the *NRC Handelsblad* website also emails a weekly report, but emphasizes the choices of individual editors: ‘All editors have access to the data of visitor-numbers, and some editors think it’s a nice little tool, but they all make their own choices’.

It is plausible that such differences in communication between print and online newsrooms and management of audience clicks information affect the ways in which this information is shared and used. Having audience metrics open on display, as is the case for *De Telegraaf*, stimulates an environment in which monitoring and sharing audience clicks are accepted as a normal routine. Effectively, it removes the decision of whether audience clicks should be monitored and shared from the hands of individual journalists, thereby challenging the stickiness of professional norms through top-down management. If, instead, audience clicks are only made available through an occasional email or non-committal access to a tool, journalists considering to use audience clicks might be less quick to compromise their professional norms individually.

Second, we look at differences in how organizations talk about the role of audience clicks in the news selection process. As popular newspapers, *De Telegraaf* and *Algemeen Dagblad* are more aimed at entertaining audiences. It makes sense that the professional judgment of the journalists on the importance of an article is less important for this role. Our interviewee from *Algemeen Dagblad* was rather candid about this priority: ‘We would rather have an article that scores, but is not that important, than an article that is important but does not score’.

In contrast, the quality newspapers explicitly distance themselves from this role conception. When we asked the editor of *de Volkskrant* whether it is important for them to cover the daily talk-of-the-town, he answered, ‘No. Then we would become the AD [*Algemeen Dagblad*]. [...] They write what the reader wants [...] We determine what we consider to be interesting’. Similarly, the editor of the *NRC Handelsblad* website emphasized that importance outweighs entertainment value: ‘News is either important or not important [...]. We have strict selection criteria. That is mainly about entertainment news. We ignore that for a large part’. The statements of the quality newspapers more clearly echo the ideals of the Trustee Model.

Still, we did find effects of audience clicks on the news selection choices of *de Volkskrant* and, to some extent, *NRC Handelsblad*. There are some editor statements that support this. The editor from *de Volkskrant* noted that while the primary criterion for news selection is their professional judgment, it is also relevant whether the story is ‘on top of mind’ of the audience: ‘If [an article] suddenly goes through the roof, then you will try to find reactions to it. It is of interest to so many people, so it is your task to inform them about it’. That there was no effect of audience clicks on the website of *NRC Handelsblad* might be because they do not share this view, but here it is also important

to consider the differences in format. As seen in Table 1, *NRC Handelsblad* produces relatively few articles. Where most newspaper websites take advantage of the online format to publish a wider range of stories, *NRC Handelsblad* makes a more specific selection in which the primary criterion of professional judgment might have more weight.

The only newspaper in our analysis that had no influence from audience clicks for both print and online news was *Trouw*. It is noteworthy that both editors from *Trouw* also displayed clear normative concerns about the influence of audience clicks. The print editor expressed this most vividly:

120 people died for Trouw in the [Second World] War and we are not going to write just anything. We do not make the newspaper that sells best, but the one we think we should make. We have a foundation that safeguards our identity. We are very aware of that identity.

We have observed that there are several differences between news organizations in the way they monitor audience clicks and how they think audience clicks should be used for news selection. In particular, we see a divide between popular and quality newspapers. These differences correspond to some extent to the findings of our quantitative analysis. There was an effect of audience clicks for the print and online editions of both popular newspapers, although the effect was relatively weak for *De Telegraaf*. For the quality newspapers, an effect was not always found. In particular, no effects were found for *Trouw*, which was also most vocal about the importance of professional norms. Overall, this reflects that the availability and top-down management of tools for the monitoring of audience clicks promote an influence of audience clicks, and a strong traditional role conception suppresses this influence.

Conclusion

We addressed the question of whether journalists use audience clicks (i.e. page views for news articles) to cater news selection to the preferences of the audience. Based on a quantitative analysis of the five largest national newspapers of the Netherlands, we conclude that for most newspapers in our study, audience clicks indeed affect news selection choices (H1). Storylines of online news articles that were in the top 5 of most-viewed articles were more likely to receive more attention in subsequent reporting, both in the print version (four out of five newspapers) and website (three out of five newspapers) of the newspaper. This result provides new quantitative support for the hypothesis that the rise of online audience metrics increases the influence of audience preferences on the work of journalists (Anderson, 2011; Lee et al., 2014; Singer, 2011; Vu, 2014).

From interviews with editors, we concluded that all newspapers in our study use tools to monitor audience clicks. This means that the main criterion for a causal relation is fulfilled: journalists are aware of audience clicks. However, the influence of audience clicks on news selection was denied or strongly nuanced. Based on our quantitative findings, we conclude that there is a discrepancy between what journalists say about the influence of audience clicks on news selection and the actual influence – they are either unaware of this influence or unwilling to admit it.

This discrepancy can be explained as a manifestation of the struggle journalists are engaged in to balance audience preferences with their professional judgment (Anderson, 2011; Lee et al., 2014; Usher, 2013). Journalists are reluctant to compromise their traditional role as gatekeepers (Deuze, 2008; Singer, 2011). Based on this traditional role conception, the use of audience clicks to cater to the preferences of the audience appears to be considered a taboo. However, due to decreasing sales and advertising incomes, and increasing competition, news organizations have shifted more toward a Market Model (Hamilton, 2004; Ryfe, 2012). This encourages journalists to cater to the preferences of the audience, and the rise of online audience metrics has made it much cheaper and easier to monitor these. Consequently, what journalists are economically encouraged to do, and what they are normatively inclined to do, are in conflict with each other.

It is not unlikely that the use of audience clicks will become more accepted as a sound journalistic practice. An interesting direction for future research would be to study whether and when this will happen, and whether breaking this normative barrier will also inflate the influence of audience clicks on the news selection process. Based on the theory that news media act as an institution (Cook, 2005), we can hypothesize that such a normative transformation requires a collective shift – coordinated or co-evolving across news organizations. Although limited in scope, our results show that a collective shift is not yet taking place in the Netherlands; journalism as an institution is still struggling to define clear professional norms for the use of audience clicks and at present sticks – at least in words – to traditional norms.

The stakes of this struggle are high, both for society and for the future of news organizations. Holding tightly to the Trustee Model might keep journalists in charge of news selection and thereby protect journalism from market logic, but at the risk of further disconnecting with their audiences and ultimately not being able to cover costs. Alternatively, they can shift more toward a Market Model and dedicate themselves to satisfying the demands of the audience. For society, the danger of this practice is that it could be at the expense of important news that is less popular. For the news organization, there is also a risk. In the long term, sacrificing their professional judgment can rob journalists of their added value, especially now that interesting information is available for free on the Internet. Our interview with the editor of the website of *Trouw* revealed that some journalists are aware of this threat. According to him, there might only be a future for quality newspapers whose professional judgment adds something to what people are already talking about.

Additionally, the role of online audience metrics in news selection warrants more research, covering more news organizations, contexts, and extending into the future. The mixed-method approach used in this study is suitable for this purpose, but certain limitations need to be taken into account, and we learned of several potential improvements. A cross-lagged regression analysis is sensitive to spurious effects. In particular, it is difficult to control for the extent to which the news selection criteria of journalists are already in sync with the preferences of the audience. Additionally, there is no single best method for determining what model to use, and different models offer different perspectives on the data. A particularly influential decision in time-series analysis is the specification of lag, so our results need to be interpreted in the context of our choices.

A notable limitation of our model is the use of days as time units. Our results show that the influence of audience clicks on news selection can last longer than a day, possibly even a week, which makes sense because journalists do not always directly have new articles available when they monitor that a storyline receives much attention from the audience. However, there can also be influences of audience clicks on journalistic practice within the day. Lee et al. (2014) found this short-term influence for news placement, and it could also be true for news selection. More research is needed on these short-term effects.

In addition, the model could be improved by taking the prominence (e.g. placement, visual cues) of online news articles into account. Prominent storylines are more likely to be viewed by the audience and also more likely to be followed up on given that it has already been judged to be worthy of prominent placement. Consequently, a portion of the relation between audience clicks on news selection could be explained by prominence. If aspects of prominence are measured, they can be taken into the equation to study its effects on audience clicks and follow-up coverage, and control for these effects.

Where the quantitative approach provides insight into the effects of audience clicks on journalistic practice, interviews with editors provide insight into journalistic norms and routines. Since the results of interviews are based on self-reports, they are not necessarily true or accurate. However, in combination with the quantitative approach, this allows us to study the discrepancy between what journalists say they do and what they actually do. Surveys are a viable alternative to interviews to get a more quantitative measurement of self-reports (e.g. Vu, 2014). This also makes it easier to include more news organizations and more editors from each organization and to get repeated measurements over time. Ultimately, longitudinal data on the ways in which different news organizations monitor and use online audience metrics are obtainable, highly relevant, and provide a rare occasion to study how journalism is evolving as an institution.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Notes

1. Digital copies of the print editions are available online for subscribers.
2. There can be more than five top 5 articles per day if there are changes in the top 5 within 1 day and less than five if articles stay in the top 5 longer than 1 day.
3. The Latent Dirichlet Allocation (LDA) topics mainly capture broad storylines that cover many news articles. To include several more specific storylines in our analysis, we also distinguished between storylines within topics based on unique key elements, such as specific people or events. For example, one topic represented news about the royal family. A large part of this news was about the ascension of King Willem-Alexander, but there was also a considerable amount of news about musicians competing to write a song for the new King. These are distinct storylines that can be of interest to journalists and audience members for different reasons.
4. Except for *NRC Handelsblad*, which publishes in the afternoon.
5. To test this, we compared models with lagged scores between 1 and 7 days (for *only* 1 day, for *only* 2 days, etc.). The model fit indeed decreases steadily as the lag increases – based on the Bayesian information criterion (BIC).

6. A common rate of decay in agenda-setting research is 0.693, which equals a half-life of 1 day (Jasperson et al., 1998). However, based on model comparisons we found that stronger effects of audience clicks are found with a less steep decay rate of 0.5. Substantially, this indicates that influence of audience clicks on news selection lasts longer than the consensus half-life used in agenda-setting research.
7. Dummy variables are used to include different media in the same model, and interaction effects with these variables are used to get effects per medium. The intercept and main effects were not included in the model so that no reference category was required.
8. The direction and significance of effects remained the same for model 1 (print) and model 2 (online), and effect sizes change only slightly. For model 3 (audience clicks), the significance of two effects changed due to minor changes in p-values. Since these differences have no implications for our conclusions, we report the model including *Algemeen Dagblad*.

References

- Anderson C (2011) Between creative and quantified audiences: Web metrics and changing patterns of newswork in local us newsrooms. *Journalism* 12(5): 550–566.
- Anderson PJ and Ward G (2007) *The Future of Journalism in the Advanced Democracies*. Aldershot: Ashgate Publishing, Ltd.
- Bakker P and Scholten O (2013) *Communicatiekaart van Nederland. Overzicht van media en communicatie (negende geheel herziene druk)*. Amsterdam: Kluwer.
- Blei DM, Ng AY and Jordan MI (2003) Latent Dirichlet allocation. *Journal of Machine Learning Research* 3: 993–1022.
- Boczkowski PJ and De Santos M (2007) When more media equals less news: Patterns of content homogenization in Argentina's leading print and online newspapers. *Political Communication* 24(2): 167–180.
- Brants K and Van Praag P (2006) Signs of media logic: Half a century of political communications in the Netherlands. *Javnost/ThePublic* 13(1): 25–40.
- Cook T (2005) *Governing with the News: The News Media as A Political Institution*, 2nd edn. Chicago, IL: University of Chicago press.
- Dahl R (1998) *On Democracy*. New Haven, CT: Yale University Press.
- Deuze M (2008) The changing context of news work: Liquid journalism for a monitorial citizenry. *International Journal of Communication* 2: 848–865.
- Fancher M (1987) The metamorphosis of the newspaper editor. *Gannett Center Journal* 1(1): 69–80.
- Galtung J and Ruge MH (1965) The structure of foreign news: The presentation of the Congo, Cuba and Cyprus crises in four Norwegian newspapers. *Journal of Peace Research* 2: 64–91.
- Gans H (1979) *Deciding What's News*. New York: Vintage Books.
- Granger CW (1988) Some recent development in a concept of causality. *Journal of Econometrics* 39(1): 199–211.
- Hallin DC and Mancini P (2004) *Comparing Media Systems: Three Models of Media and Politics*. Cambridge: Cambridge University Press.
- Hamilton J (2004) *All the News that's Fit to Sell: How the Market Transforms Information into News*. Princeton, NJ: Princeton University Press.
- Hollanders D and Vliegenthart R (2008) Telling what yesterday's news might be tomorrow: Modeling media dynamics. *Communications* 33(1): 47–68.
- Jasperson AE, Shah DV, Watts M, et al. (1998) Framing and the public agenda: Media effects on the importance of the federal budget deficit. *Political Communication* 15(2): 205–224.

- Lee AM, Lewis SC and Powers M (2014) Audience clicks and news placement: A study of time-lagged influence in online journalism. *Communication Research* 31(4): 505–530.
- McManus J (1995) A market-based model of news production. *Communication Theory* 5(4): 301–338.
- Paulussen S, Heinonen A, Domingo D, et al. (2007) Doing it together: Citizen participation in the professional news making process. *Observatorio (OBS*)* 1(3): 131–154.
- Ryfe DM (2006) Guest editor's introduction: New institutionalism and the news. *Political Communication* 23(2): 135–144.
- Ryfe DM (2012) *Can Journalism Survive: An Inside Look at American Newsrooms*. Cambridge: Polity Press.
- Schudson M (2003) *The Sociology of News*. New York: W. W. Norton & Company.
- Shoemaker PJ (1991) *Gatekeeping*. Newbury Park, CA: SAGE.
- Shoemaker PJ and Vos T (2009) *Gatekeeping Theory*. New York: Routledge.
- Singer JB (2011) Community service: Editor pride and user preference on local newspaper websites. *Journalism Practice* 5(6): 623–642.
- Usher N (2013) Al Jazeera English online: Understanding web metrics and news production when a quantified audience is not a commodified audience. *Digital Journalism* 1(3): 335–351.
- Vliegthart R and Walgrave S (2008) The contingency of intermedia agenda-setting: A longitudinal study in Belgium. *Journalism & Mass Communication Quarterly* 85(4): 860–877.
- Vu HT (2014) The online audience as gatekeeper: The influence of reader metrics on news editorial selection. *Journalism* 15(8): 1094–1110.
- White DM (1950) The gatekeeper: A case study in the selection of news. *Journalism Quarterly* 27(4): 383–390.

Author biographies

Kasper Welbers is a PhD student at the Department of Communication Science at the VU University Amsterdam. His research focuses on how the gatekeeping process has changed due to the proliferation of new media technologies.

Wouter van Atteveldt is Assistant Professor in the Department of Communication Science at the VU University Amsterdam. He studies methods for automated content analysis and is the principal developer of the Amsterdam Content Analysis Toolkit (AMCAT, amcat.vu.nl). He is currently pursuing research – funded by the Dutch National Science Foundation (NWO-VENI) – into the role of media statements by politicians in the relation between press and politics.

Jan Kleinnijenhuis is Professor of Communication Science at the Department of Communication Science of the VU University Amsterdam. His research deals with political and economic journalism, with the emphasis on news selection and news effects.

Nel Ruigrok founded her own research company, LJS Media Research, which conducts media analyses for a number of (commercial) organizations. She also offers lectures and courses at national and international universities. Research interests concern political communication in general and especially the role of media during election campaigns, and the role of the media during conflicts.

Joep Schaper is a PhD Candidate in the research group Political Economy and Transnational Governance (PETGOV) in the Department of Political Science of the University of Amsterdam. His PhD project 'Causes of Immigration Policy Restrictiveness' addresses the question what causes countries to open or close their borders to different kinds of immigrant groups.