



UvA-DARE (Digital Academic Repository)

Political relevance in the eye of the beholder: Determining the substantiveness of TV shows and political debates with Twitter data

Boukes, M.; Trilling, D.

DOI

[10.5210/fm.v22i4.7031](https://doi.org/10.5210/fm.v22i4.7031)

Publication date

2017

Document Version

Final published version

Published in

First Monday

License

CC BY-SA

[Link to publication](#)

Citation for published version (APA):

Boukes, M., & Trilling, D. (2017). Political relevance in the eye of the beholder: Determining the substantiveness of TV shows and political debates with Twitter data. *First Monday*, 22(4). <https://doi.org/10.5210/fm.v22i4.7031>

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (<https://dare.uva.nl>)



- [HOME](#)
- [ABOUT](#)
- [LOGIN](#)
- [REGISTER](#)
- [SEARCH](#)
- [CURRENT](#)
- [ARCHIVES](#)
- [ANNOUNCEMENTS](#)
- [SUBMISSIONS](#)

[OPEN JOURNAL SYSTEMS](#)

[Journal Help](#)

Home > Volume 22, Number 4 - 3 April 2017 > **Boukes**



Political relevance in the eye of the beholder: Determining substantiveness of TV shows political debates with Twitter by Mark Boukes and Damia

Abstract

Addressing the call to move beyond a simple genre classification of TV shows as either substantive (hard) news or non-substantive (soft) infotainment, we propose using social media reactions to determine a program's political relevance. Such an approach provides information that goes beyond genre or content characteristics and reflects what really reaches an audience. Analyzing tweets about two Dutch talk shows and four U.S. primary debates, we show that audience responses to television programs differ considerably regarding their political relevance. Thereby, we demonstrate how examining online audience reactions can be employed as a sophisticated and valid way to assess the political relevance of TV programs.

Contents

- [Introduction](#)
- [Theoretical background and related research](#)
- [Method](#)
- [Results](#)
- [Discussion and conclusion](#)

Introduction

Comparing political media today and some decades ago, it is hard to deny that there have been dramatic changes. In this paper, we focus on two trends that have become more prominent and are still evolving: (a) the softening of political news and (b) the use of social media as a second screen while watching political programming to express one's thoughts and

USER

Username

Password

Remember me

JOURNAL CONTENT

Search

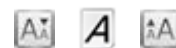
Search Scope

All

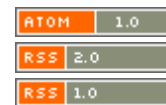
Browse

- [By Issue](#)
- [By Author](#)
- [By Title](#)
- [Other Journals](#)

FONT SIZE



CURRENT ISSUE



ARTICLE TOOLS

- [Abstract](#)
- [Print this article](#)
- [Indexing metadata](#)
- [How to cite item](#)
- [Email this article \(Login required\)](#)
- [Email the author \(Login required\)](#)

feelings. We demonstrate how data from this second screen may be employed to analyze developments towards less substantive political coverage in traditional media.

Focusing on the first development, “lighter” formats have been introduced to convey political topics. While citizens have massively tuned out some traditional news programming (Mindich, 2005) and moved from public to commercials broadcasters (Bakker and Scholten, 2011), television programs that combine political information with entertainment have become increasingly popular (Williams and Delli Carpini, 2011). The question, yet, remains *how much* political information such infotainment provides. Infotainment, a neologism of “information” and “entertainment,” implements characteristics of both news and entertainment programs in one outlet (Otto, *et al.*, 2016); yet, how much substantive information it contains is difficult to assess.

Just as infotainment does, political debates do not only provide substantive content but are for a large part driven by non-substantive format characteristics, such as personal attacks (Racine Group, 2002) or hypes (Fox, *et al.*, 2005). Political debates and infotainment may thus be watched to acquire political information but also out of hedonic motivations — simply because they are entertaining (see, *e.g.*, Roth, *et al.*, 2014). The question, however, remains *how* politically substantive these programs actually are themselves, but especially *vis-à-vis* one another. Infotainment and political debates, thus, share that both (only) partly consist of politically substantive information, and this paper develops an approach to determine their relative degree of political substantiveness.

The second development we look at is the advent of social media, which has considerably changed how citizens communicate with each other, but also with political actors. Without constraints in time and space, people can now (at least in theory) share and discuss political ideas. And, they indeed do so: For example, the question how people use a so-called *second screen* to comment on political debates has been studied repeatedly in recent years (*e.g.*, Jungherr, 2015; Shamma, *et al.*, 2009; Trilling, 2015; Vergeer and Franses, 2016). In line with Shamma and colleagues, who argue that tweeting about television programs can be described as “community annotation” because the tweets reflect the structure of the program people are watching (Diakopoulos and Shamma, 2010; Shamma, *et al.*, 2010; Kennedy, and Churchill, 2009), we use such social media reactions to TV programs and political debates to assess *how* politically substantive these programs are.

Such insights are important, because scholars of public opinion decide which outlets to include or exclude from their studies (*i.e.*, content analyses or surveys) depending on whether these would be political or not. Whereas for newspapers or newscasts the political relevance is not doubted, entertainment programs, such as talk shows, still do not get the academic scrutiny they deserve (Williams and Delli Carpini, 2011). Developing a method to demonstrate the substantiveness of these outlets is of crucial importance for breaking down the artificial hierarchy that exists between news programs and infotainment (Boukes and Boomgaarden, 2015). This methodology will also provide practitioners and policy-makers with a better way to assess public perceptions, content and political relevance of programs. Moreover, it will raise public awareness of the role that infotainment may play in the public sphere of politics.

We collected tweets about two Dutch talk shows and U.S. primary debates to answer the following research question: *How can social media data be used to assess the relative degree of substantive political information that is transferred to the viewer through infotainment or political debates?* The content of tweets is examined to see how strongly a particular program triggers substantive political reactions as a way to determine its political relevance. Thereby, we show how scholars can move beyond the so-frequently criticized reliance on genre to decide on a programs’ political relevance (*e.g.*, Lehman-Wilzig and Seletzky, 2010), and instead could employ empirical data to bolster the arguments about the substantiveness of the programs in question.

The rationale behind this method is that using formal characteristics of a program like those found in its description is not sufficient to infer whether it is hard or soft news, informative or pure entertainment, substantive or superficial, when we want to say something about the consequences of watching these programs. After all, media effect theories refer to the content that message recipients in one way or the other *process*. Just relying on genre classifications or content analyses, thus, does not

ABOUT THE AUTHORS

Mark Boukes
University of
Amsterdam
Netherlands

.....
Damian Trilling
University of
Amsterdam

approach the core of what is at the root of any media effect theory: the processing of media content. Instead, we plead for the use of audience responses, by analyzing discussions on social media, to make more sophisticated claims about the political substantiveness of specific programs.

Audience reactions reveal what viewers actually pick up from shows that they tweet about. This is arguably much more important than a genre classification and goes beyond an analysis of content characteristics. The substantive information that politicians express in political debates, for example, has a weaker effect on viewers than the way politicians present themselves (Racine Group, 2002): Presentational style, thus, seems to override specific content in terms of effects. So, apparently it is not only about what topics have been discussed and are *presented* to the audience — what a content analysis could reveal — but mainly about what attracts the *attention* of the audience, which eventually has an effect. An analysis of audience responses provides a deeper insight into what actually triggers the attention of viewers, and may thus have an effect. Audience reactions show what actually came across and, therefore, are a more detailed indicator of an outlet's political relevance and potential influence.

Research has shown that a non-negligible amount of tweets on political TV programs address politically substantive issues (D'heer and Verdegem, 2014; Jungherr, 2015; Trilling, 2015). Hence, it will be possible to distinguish between programs that evoke a higher or lower proportion of substantive reactions. Using techniques of word-frequency based corpus comparison, we aim to investigate the topics that are discussed regarding different TV programs, which subsequently are used as an indicator of a program's political relevance. The implications of this are twofold: First, it contributes to the knowledge about the potential outcomes of and public reactions towards television programming; and second, it advances an understanding of how social media data can be used to study political TV shows.



Theoretical background and related research

The theoretical framework for this study starts off with explaining the problems and difficulties of relying on program formats or content analysis to determine the political relevance of television programs — or any kind of media content. Alternatives are discussed, and a claim is made why looking at audience responses is a valid way of evaluating whether a program is politically relevant or not. Subsequently, the potential of analyzing social media data for understanding political behavior is elaborated. Special attention is given to the use of social media as a second screen while watching political programming.

Spectacle or substance? Disentangling the political relevance of TV programs

Studies on soft news and infotainment very often make a rough division between soft news and hard news outlets (*e.g.*, Baum, 2003; Rittenberg, *et al.*, 2012). Reinemann, *et al.* (2012) distinguish three dimensions that separate hard from soft news: (1) news is politically relevant or focuses on societal conflicts (yes: hard; no: soft); (2) whether news focuses on society, politics or broader trends (hard) versus a focus on individual cases by means of episodic frames (soft); and, (3) whether the style in which news is presented is impersonal and unemotional (hard) or personal and emotional (soft). However, dividing all news into either being hard or soft is problematic because most programs are not purely hard or soft, entertaining or informing. Such thinking in terms of “either-this-or-that” extremes neglects the observations that most news coverage reflects characteristics that apply to both hard and soft news (see Boukes and Boomgaarden, 2015; Zelizer, 2007). Categorizing programs in these two mutually exclusive alternatives is, therefore, artificial.

Another point of critique is the reliance on genre to determine whether a program is hard or soft news. After all, increased market pressures caused hard news producers to also be aware of its business responsibility to attract bigger audiences and sell commercials (McManus, 1995). Hence, traditional news programs increasingly report on political and serious issues while using entertainment techniques (*e.g.*, Vettehen, *et al.*, 2011). To say that traditional news programs should all be labeled “hard” and substantive while infotainment programs would by definition be “soft” and non-

substantive, thus, seems too simplistic. The content of infotainment programming is in certain cases equally issue-oriented and informative as news programming (e.g., Fox, *et al.*, 2007; Haigh and Heresco, 2010; Matthes and Rauchfleisch, 2013).

In a nutshell, claiming that infotainment programs are less politically relevant than traditional news and, therefore, should be classified as soft news *per se*, is too harsh claim and takes an elitist perspective on media content (Althaus, 2012). Such a classification, thus, is flawed and should be replaced by other methods. One alternative is to look at audience characteristics and how these overlap between programs (Boukes and Boomgaarden, 2015). However, this requires extensive survey data that often is not available and expensive to collect. It would be helpful to determine programs' political relevance based on data that are more readily available and simultaneously reflect content characteristics of these programs.

The same issue, whether programs are substantive or merely dealing with non-political issues, is surrounding political debates (Fox, *et al.*, 2005). Although previous studies have shown that debate content is most of the time about issues and provides substantive information (Racine Group, 2002), candidate statements made in televised debates are generally a combination of *substantive* political remarks regarding past political records or future political plans with *non-substantive* personal attacks on one's competitors or a defense of one's (not-so-flawed) character (de Nooy and Maier, 2015). Moreover, meta-analysis found that debates generally affect perceptions of the candidates' personalities, but they do not have much of an effect on the perceptions of their competence (Benoit, *et al.*, 2003). The provided content — mainly substantive — may thus not be the content that is most prominent in the eyes of the audience.

The substantiveness of political debates is often subject of popular debate. One of the most memorable debate expressions is Ronald Reagan's non-substantive "*There you go again*" against President Carter. This obviously rehearsed statement simply intends to disarm his opponent rather than bringing forward his own policy plans. Regarding the 2016 primary debates for the Republican and Democrat parties, the media paid significant attention to the substantiveness of debates with headlines, such as "*Democrats see a more substantive, if sleepy, debate than rowdy GOP show*" (Rucker, 2015). A clear distinction seemingly is made between debates that, on the one hand, are substantive and talking about political topics, such as immigration, tax proposals or education (Martin and Healy, 2015) and those, on the other hand, in which it is mostly about raised voices, name calling, clashing, blaming, accusing the other of being a liar, and ridiculing one another (Sullivan, 2016).

While journalists use their own experiences and observations to assess the political relevance of debates, scholars still lack methods of doing this in a valid, structural and replicable manner. In this paper, we propose to analyze audience reactions to assess the political substance of television programs. The reactions of an audience can reveal the *actual contribution* of media content over and beyond what it promises to do (reliance on genre classification) or what it seems to do (reliance on content analysis). After all, central in much thinking about how political news would ideally look like from a democratic perspective is its ability to mobilize citizens' engagement in public life by means of discussing politically relevant matters (Ferree, *et al.*, 2002). This is a feature of political media content that is too often neglected when assessing its democratic qualities (Althaus, 2012), while it is a good indicator of its political relevance.

The second screen as remote co-viewing, place for debate and community annotation

The use of social media while watching television has recently attracted scholarly attention, using terms such as "social TV," "second-screen usage," or "remote co-viewing." Although this has not yet led to the emergence of one well-defined theoretical framework to study such phenomena, existing research from different disciplines provides us with a useful starting point for understanding the use of social media while watching TV.

According to Pittman and Tefertiller (2015), "[s]ocial media based, second-screen co-viewing is a key indicator of involvement with television." Consequently, television stations actively promote this behavior to enhance audience engagement and audience loyalty. Because online commenting on television programs while watching it has increasingly become popular (e.g., Courtois and D'heer, 2012), using audience responses to classify television programs is now possible.

While sometimes, specific apps for tablets and smartphones are used for commenting on the so-called “second screen,” existing social media sites like Facebook and Twitter are more frequently used. In the case of Twitter, the openness of the platform encourages audience participation and does not require central action, such as the creation of a specific group or platform, allowing everyone to respond to TV content immediately.

Responding on social media while watching a program is a form of (remote) co-viewing (Buschow, *et al.*, 2014). It is known that co-viewing (*i.e.*, watching TV together) is an important factor people consider in their choice for TV programs, not only for entertainment but also for news (Wonneberger, *et al.*, 2011). Assuming that watching TV is a social experience and that people enjoy talking about the programs they are watching, it makes sense to assume that these activities are not only occurring face-to-face, but on social media too.

Research has therefore addressed second-screen communication under the name of “social TV” (Buschow, *et al.*, 2014). Doughty, *et al.* (2012) describe this as “inviting someone to ‘share your sofa’ through tweet mentions and retweets.” [1] This is also reflected in the motivations of viewers to engage in this kind of behavior. Studying tweets on a variety of German television programs (including both entertainment and informative ones), Buschow, *et al.* (2014) conclude that “tweets seem to be motivated by two desires: to interact with the community and/or to engage with the program.” [2] Yet, the latter was most prevalent, and indicates that social media reactions reflect to some extent the content of the programs that people are watching.

While studying second-screen usage from a co-viewing perspective is useful if one is interested in social-psychological processes and the motives of users to engage in such behavior, another angle is to focus on the added value for public debate. Relying on the notion of a public sphere, one can argue that tweeting about a program can be a platform for the exchange of arguments. If it is true that at least a fair share of those who tweet about a program do so in order to contribute to a political debate, then a sufficiently large amount of their tweets can reasonably be expected to contain substantial information that is passed on from the program. And indeed, zooming in on the most retweeted tweets during a political debate, Jungherr [3] estimates that 25 percent of these high-impact tweets comment on the debate itself.

In general, tweets about political debates reflect the topics of the debate, but can also add or emphasize a large share of other aspects, both politically relevant and irrelevant ones (Shamma, *et al.*, 2009; Trilling, 2015; Vergeer and Franses, 2016) [4]. So, because a vast majority of social media reactions refers to the content or the guests of political television programs (D’heer and Verdegem, 2014), these reactions may be employed to determine the political substantiveness of these shows through the eyes of the audience. And indeed, it has been shown that second-screen tweeting can be understood as a form of “community annotation” of television programs, because both the structure and the content of the program are well reflected in tweets (Diakopoulos and Shamma, 2010; Shamma, *et al.*, 2010, 2009). More specifically, Vergeer and Franses (2016) show that topics like “housing”, “care for the needy,” and “Europe” align very well when comparing the content of TV debates and Twitter comments (see also Yldrm, *et al.*, 2016).

One might have objections regarding the (un)representativeness of the user base on Twitter. This is indeed true if one makes inferences about the amount of support for a certain specific policy or candidate, and this is why attempts to use Twitter analysis as a replacement for polling have been extensively criticized (Gayo-Avello, 2013; Jungherr, *et al.*, 2011). Yet, this is less of a concern in our case as we are not interested in obtaining *exact* estimates of percentages of the population that hold a given opinion. Instead, we follow Jungherr, *et al.* (2016), who argue that tweets are a good indicator of *relative* attention.

Because those who online post content about politics are different from those who do not (*e.g.*, more interested in politics, overrepresentation of males; Bakker, 2013), they cannot be treated as a random sample of the whole population. The same is true when the population is restricted to audiences of TV programs: It is not a random sample of television watchers that tweet about programs. Therefore, we cannot infer estimates of precise *percentages* of how important a given issue is to an audience. However, there is little reason to assume that the *set of issues* that the tweeting subset of the audience picks up from one program differs substantially from

the issues that the non-tweeting audience picks up. Thus, while being only of limited use for obtaining exact estimates of proportions, we argue that social media data are a sufficiently valid proxy for securing a picture of the issues that people talk and care about while watching one TV program relative to another program. In particular, even despite the bias introduced by the non-random sampling, it allows us to obtain a picture of the *relative* attention that is paid to issues between programs: After all, the decision to tweet versus not to tweet most likely mainly depends on the characteristics of the user (*i.e.*, what strikes him or her or is salient), and only to a very limited extent which of two (similar) programs the user is watching.

Specific research questions

So far, we have argued that analyzing audience responses potentially is a viable way to measure the substantiveness of TV programs. We have also shown that social media data, and tweets in particular, are a useful source for information about audience responses. In two case studies, we demonstrate how the most characteristic words of two similar TV programs can be used to determine their political relevance. We compare two Dutch infotainment talk show programs (*Pauw* and *DWDD*) in the first case study, and for the second case study we focus on political debates for the U.S. 2016 primary elections between respectively Democratic and Republican candidates. To answer the overarching research question posed in the introduction, we specifically ask:

RQ1: How politically substantive is DWDD versus Pauw according to tweets sent about these two infotainment talk shows?

RQ2: How politically substantive were the primary Republican and Democratic debates according to the tweets sent about these debates?



Method

Case study 1: DWDD and Pauw

Like many European countries, the Netherlands has a strong public service broadcasting system. While private stations play an important role, most news and current affairs programs are watched on one of the three public stations. We focus on two infotainment talk shows, *DWDD* (“De Wereld Draait Door” [The world keeps turning]) and *Pauw*, which are related to news and current affairs, but also deal with entertainment regularly. Both programs are broadcasted every weekday on the public service station *NPO 1* and are produced by the same broadcaster *VARA*. Both programs are equally popular, as the average Dutch adult watches *DWDD* 2.0 and *Pauw* 1.9 days per week (Bos, *et al.*, 2014).

Both shows have one middle-aged, male host (*DWDD*: Matthijs van Nieuwkerk, *Pauw*: Jeroen Pauw) and invite a number of guests (approximately four) with whom current issues are discussed for about 10 minutes. Both programs last about 50 minutes, and are produced in the same building in Amsterdam. One formal difference is the timeslot: 19.00 to 19.50 (*DWDD*) versus 23.05 to 23.55 (*Pauw*).

In the Netherlands, eight percent of the population use Twitter weekly ‘for news’ (Reuters Institute for the Study of Journalism, 2016). While there seems to be an overrepresentation of younger people; recently, younger people seem to drop out and older people seem to join (*e.g.*, Bakker, 2013; Oosterveer, 2014; Newcom Research, 2016).

Case study 2: U.S. primary debates

Between 1 February and 14 June 2016, a series of primary elections and caucuses, in the run-up to the U.S. presidential elections on 8 November 2016, were held. These primary elections on who will be the party’s candidate for the next presidential election are accompanied by several televised debates. We study reactions to two Democratic and two Republican televised debates: the New Hampshire Democratic Debate (4 February 2016), New Hampshire Republican Debate (6 February), Wisconsin Democratic Debate (11 February) and the South Carolina

Republican Debate (13 February).

Data collection and preprocessing

We installed *DMI-TCAT* (Borra and Rieder, 2014) on a virtual machine on a cloud computing platform. This software continuously queries the Twitter streaming API to collect tweets containing specific keywords. In the Dutch case, the hashtags #dwdd and #pauw are actively promoted by the two television shows and used by viewers that wish to online comment on the programs. Between 6 January and the beginning of the programs' summer break in 2015 (*DWDD*: 24 May; *Pauw*: 29 May), we collected those tweets that included the hashtags #dwdd and #pauw. This encompasses half a television season (20 weeks), and most likely is a rather generalizable time frame, because it did not include major political events, such as national elections [5].

Because we are only interested in tweets that were sent while watching the programs, a *Python* script was written to remove all tweets that were sent on weekends (as the programs aired only on weekdays) or outside of the broadcasting time. We also removed all days on which less than 500 tweets were sent, to filter out occasions when the show was not broadcasted (e.g., during holidays). The dataset constructed in this way contains $N_{DWDD} = 78,162$ and $N_{PAUW} = 56,404$ tweets.

The same *Python* script was used to do some basic pre-processing, especially removing the hashtags themselves, links, usernames, numbers and stopwords (i.e., extremely frequent words without clearly interpretable meaning like 'a', 'the', 'her', 'his'). Terms that directly refer to the programs (like variations of their names and the names of the hosts) were added to this list of stopwords.

For Case Study 2, we collected tweets that used the hashtags #gopdebate or #DemDebate, which were the "official" hashtags promoted by the broadcasters and consistently used by those who commented on the debates. We retrieved $N_{Rep} = 243,983$ and $N_{Dem} = 242,848$ tweets, which were sent between the first and the last minute of the debates. All preprocessing steps were analogues to those in Study 1.

Data analysis

The pre-processed data were analyzed by a self-written *Python* script. We determined the frequencies of most-occurring words in the corpora of *DWDD* and *Pauw* tweets for the first case study, and in the corpus of the Republican debate versus the corpus of the Democratic debate for the second case. Rather than analyzing single tweets, we compare two times two corpora, each of which consists of all concatenated tweets about the program.

To identify the most typical terms for each of the two corpora (e.g., Corpus 1 being all tweets about *DWDD*, Corpus 2 being all tweets about *Pauw*), we calculated the log-likelihood for each word to occur in one of these two corpora [6]. This value shows how well the occurrence of a given word indicates whether a text belongs to one of the corpora (i.e., *DWDD* versus *Pauw*). In other words, if a term occurs much more (or less) frequently than expected by chance in one of the two corpora, the value of the log-likelihood increases. The higher the log-likelihood value for a given term, the more characteristic this term is for one of the two corpora.

We followed Rayson and Garside's (2000) formula to calculate this log-likelihood: Given the frequency a of a word in Corpus 1 (i.e., *DWDD*), its frequency b in Corpus 2 (i.e., *Pauw*), the total length in words of Corpus 1 c , and the total length in words of Corpus 2 d , the expected frequency of the word in Corpus 1 can be calculated as follows:

$$E1 = c(a+b)/(c+d)$$

and its expected frequency in Corpus 2 as:

$$E2 = d(a+b)/(c+d)$$

The log-likelihood compares the observed frequencies a and b with the expected frequencies $E1$ and $E2$. Formally, it can be calculated as:

$$LL = 2((a \ln(a/E1)) + (b \ln(b/E2)))$$

To say something about the political substantiveness of the tweets, we did not only need to calculate how typical their words were in occurring in

either the tweets about *DWDD* or *Pauw* (or Democratic vs. Republican debates), but we also had to identify these as being political or not. To do so, we used the following conceptual definition: A word is politically substantive in case it refers to a political actor, a political topic, or a topic or source usually associated with political matters. To be able to do so, the 300 words with the highest log-likelihood values were manually coded by the first author to determine whether these were (1) politically substantive or (2) politically irrelevant. Because the Dutch infotainment talk shows often addressed hard news topics that were societally relevant but not directly political, the words in tweets about *DWDD* and *Pauw* were identified as “substantive” if these referred to (a) a person who is a political actor now or in the past, (b) a political topic that has been discussed in the political arena of Parliament, or (c) a current affairs news topic, another journalistic media outlet, or a political topic expert, which are very likely to have been referred to in the context of political issues at the moment of the broadcast.

Analyzing content on the word level, on the one hand, offers a high level of precision but, on the other hand, limits the ability to take context into account. Hence, the current approach does not allow inferring whether a politically substantive word is applied in a non-political context, or vice versa. The strength of our approach, instead, lies in its simplicity: The concept of word frequencies is immediately understandable, also for stakeholders like practitioners and policy-makers, and it avoids introducing a “black box” using an opaque technique to determine context.

To ease interpretation of the results and to showcase how they could be communicated to practitioners and policy-makers, we wrote a script using the *Python* module *matplotlib* to visualize the most characteristic words in figures. We sorted all words on their LL value and selected for each corpus the top n words where the observed value was higher than the expected value. Thus, we have two columns (for each corpus one) with n words each (see Figures 1 and 2). These were plotted in a figure in the following manner: Horizontally, the words were positioned further away from the middle of the graph for higher LL values. Additionally, the base font size was multiplied by the square root of the relative frequency of the word in each corpus; hence, more frequently used words were plotted with larger font sizes. This allows distinguishing between infrequently occurring, but highly characteristic words (because the *difference* in their share between the corpora is still high) and words that occur very frequently, but are not very characteristic (because they occur in both corpora equally frequently). The former would appear far from the middle (*i.e.*, highly characteristic), but small (*i.e.*, infrequently occurring); while the latter would be closer to the middle (*i.e.*, less characteristic), but with a larger font size (*i.e.*, high frequency). Finally, the words that were coded as being substantive were marked with a shaded background: When there are more background shaded words in one corpus, this indicates that a higher proportion of the words were of politically relevant nature.



Results

Substantiveness of infotainment talks shows: DWDD versus Pauw

The comparison of the most characteristic words of the *DWDD* and *Pauw* corpora reveals striking differences. As Figure 1 shows, virtually all of the most characteristic words in the *Pauw* corpus are politically substantive. By contrast, many tweets related to *DWDD* refer to artists like Anouk or Douwe Bob, sports (“Elfstedentocht” [an ice skating event], “Ajax,” “Feyenoord” [two soccer teams]), or other TV programs. The words that were politically relevant often referred to individuals instead of topics (Aboutaleb, Mulder, Prem, Churchill, Fekiz76 — all of them being either politicians or political talking heads). As the comparably small font size indicates, despite occurring more often than expected by chance, these words also only comprised a rather small share of the words in the corpus.

Notwithstanding this general pattern, *DWDD* occasionally triggered politically relevant responses. For example, the terrorist attack on the French magazine *Charlie Hebdo* was frequently mentioned, just as the supporting statement #jesuischarlie.



Figure 1: Most characteristic words for *DWDD* (left) and *Pauw* (right). Note: The further from the 0-point on the x-axis, the higher the log-likelihood, *i.e.* the higher the probability of the word in the corpus compared to the union of both corpora. The larger the font size, the more frequently the word occurs within its corpus.

If we zoom out and look into the 300 words with the highest log-likelihood, we see our initial picture being confirmed. Out of the 172 words characteristic for *Pauw*, 25 (15 percent) were politically irrelevant, and 133 were politically relevant (77 percent): 24 words (14 percent) referred to a political topic, 41 (24 percent) to a political actor, 16 (9 percent) to a political commentator or expert, whereas 52 (30 percent) referred to a

current affairs-related topic [7]. In sharp contrast, only one (one percent) of the 128 DWDD-words was about a political topic, seven (five percent) about a political actor, eight (six percent) about a political commentator, 18 (14 percent) a current affairs topic. So in total, only 34 words (27 percent) were politically relevant, whereas 80 words (63 percent) were irrelevant. All in all, the conclusion seems to be justified that *Pauw* is more politically relevant than *DWDD*.

Substantiveness of political debates: Democrats versus Republicans

Turning to Case Study 2, as [Figure 2](#) shows, tweets about the Democratic (left) and the Republican (right) debates actually contained a similarly high share of politically substantive words. We also see that completely different topics have been addressed: The tweets about the Democratic debate — among other issues — addressed issues of campaign finances, racism and health care. Most prominent, however, is the Sanders/Clinton controversy about the candidates' relationship to Henry Kissinger, which also gained considerable media attention (*e.g.*, Allen, 2016). Additionally, Figure 2 shows that next to mentioning substantive topics, the second-screen is used to voice support for a candidate, as evident by the use of additional hashtags like #cruzcrew, #imwithher or #debatewithbernie.



Figure 2: Most characteristic words for Democratic (left) and Republican (right) debates. Note: The further from the 0-point on the x-axis, the higher the log-likelihood, *i.e.* the more characteristic the word is for the respective debate compared to the union of both corpora. The larger the

the word occurs within its corpus.

The tweets about the Republican debate, in contrast, are less diverse and highlight mainly one topic: the discussion about the political consequences of the death of U.S. Supreme Court Judge Antonin Scalia. This difference in breadth is also reflected in the share of substantive words once we expand our selection beyond the 35 words depicted in Figure 2. Taking all words from both parties' debates with a log-likelihood value of 200 or higher, 92 out of these 304 words can be classified as politically substantive, which amounts to 30 percent. Out of the 93 words from the Democratic debates, 36 were substantive (38 percent). This was considerably more than the 56 out of 211 (27 percent) words from tweets about the Republican debates. Hence, we conclude that the Republican debates were less substantive than those of the Democrats.



Discussion and conclusion

We have investigated whether audience reactions to television programs are useful to empirically assess the political substantiveness of these programs. In particular, we were interested in creating a more differentiated picture of these media outlets' relevance than what is the case with genre-based classifications like rather arbitrary distinctions between hard and soft news (see, *e.g.*, Lehman-Wilzig and Seletzky, 2010) or journalistic evaluations of substantive versus "nasty" debates (see, *e.g.*, Martin and Healy, 2015; Sullivan, 2016). Comparing tweets about two Dutch talk shows, which both would fall into the infotainment genre (*i.e.*, soft news), and tweets about U.S. primary debates between Republicans and Democrats, we demonstrated that analyzing audience responses indeed allows getting a more nuanced picture of the political relevance of these programs. Our analysis reveals that despite the great similarity in the programs' formal characteristics, audience reactions differed substantially.

Our first case study shows that categorizing both *DWDD* and *Pauw* as soft news without any further distinction does not do justice to how the audience apparently appreciates these programs differently. They clearly perceived *Pauw* to be more politically relevant than *DWDD*. This finding is in line with research on audience characteristics of these shows, because *Pauw* viewers tend to enjoy programs with news content while those of *DWDD* enjoy popular news formats more (Bos, *et al.*, 2014). Hence, we argue that it is possible to place TV programs on a scale from more to less politically relevant, and that this can be done by examining the responses of the audience on social media.

Also in our second case study, we were able to demonstrate that two programs (*i.e.*, political debates) with very similar formal characteristics evoked different audience reactions. The Democratic debates, on the one hand, evoked reactions about a diverse set of political topics, while the Republican debates, on the other hand, prompted reactions to only a relatively limited number of political topics. Hence, the conclusion can be drawn that the analyzed Democratic debates were more substantive than those of the Republicans. This empirically supports the viewpoints expressed by journalists (*e.g.*, Rucker, 2015); providing a validation of these findings. Yet, while the share of substantive words was higher in tweets about the Democratic debates, the difference was less pronounced than in our first case study comparing the Dutch infotainment talk shows.

The fact that the outcome of our analysis aligns with journalists' perceptions demonstrates the validity of our approach, which is replicable and therefore can be used for scientific purposes in contrast to the journalistic accounts. Another example for the validity of our method is the high popularity of the Kissinger topic in the tweets about the Democratic debates — indeed one of the major points that was discussed in the media after the debate (*e.g.*, Allen, 2016) signaling its relevance and prominence.

Though our approach is a rather straightforward one, it allows us to shine a different light on the political nature of television programs than has been done before. We argue that the way audiences interpret media content should be a much stronger determinant in the categorization of media outlets being politically relevant or not. After all, a program may be part of a serious news genre or contain particular content characteristics, but if the


audience does not perceive it in that way, what is such a categorization worth? Social media data, such as tweets, can be an indicator of this interpretation of the audience and are a powerful, alternative tool to investigate and categorize media outlets in terms of their political relevance.

We have demonstrated a simple-to-apply yet insightful way to use second-screen data for extracting information to assess the degree of political substantiveness of different television programs. While there are a number of other methods to extract meaning from second-screen corpora, our method offers easy to interpret insights and has the advantage that it focuses on the most *characteristic* words: As we can safely assume that noise occurs in both corpora at a roughly equal rate, this noise is not characteristic for only one of the TV programs and disappears from the graphs due to a low log-likelihood value. There are several ways of advancing our methodology. For example, one could consider using bigrams or trigrams rather than single words to take the context into account in which words are used; or one could extend our approach to compare not two, but multiple corpora, using other measures to determine the relative importance of a word, like the tf-idf measure, which weighs word frequencies. It would even be possible to apply our technique to generate real-time graphs of a debate.

Regarding the interpretation of the most characteristic words, the choice of the number of words, n_{cutoff} , that are interpreted (in this study we used $n_{cutoff} = 300$ words) is inherently somewhat arbitrary and it is difficult to specify a general rule. Choosing a too large n_{cutoff} bears the risk of including words that are actually not very much characteristic for one of the corpora any more, as they have either a low absolute frequency or a low difference in the frequencies between the two corpora. Choosing a too small n_{cutoff} , in contrast, focuses on a too small part of the corpus to provide detailed insights. The choice of n_{cutoff} , thus, should be informed by both the interpretability of the results and by the numeric values. A limitation of the current approach is that it depends on the characteristics of the specific audience that is commenting on the programs. Although political interest appears to be only a relatively weak predictor of watching news and current affairs programs in the Netherlands and that situational factors are much more important (Wonneberger, *et al.*, 2011), we can reasonably expect that *Pauw* viewers are slightly more politically interested than *DWDD* viewers. This could cause an unwanted feedback loop in our data, because the audience of *Pauw* might simply be more inclined to send political tweets. However, not only is the difference of the political interest of the two audiences rather minor (Wonneberger, *et al.*, 2011), there is also a considerable overlap between the audiences (Bos, *et al.*, 2014; Trilling and Schoenbach, 2015) [8]. Therefore, while one has to consider the possibility of a skew in the data because of different audiences, we are confident that in our case, this skew is of an acceptable magnitude.

For now, we applied our approach to two case studies; two different types of TV programs (infotainment talk shows and political debates) in two different countries (the Netherlands and the U.S.). Yet, by collecting tweets about other television programs, Web sites or other media, one can easily follow the same approach to determine to which extent these outlets are of a political nature. Hence, scholars can overcome the oversimplified dichotomy of dividing media outlets in either soft or hard news categories (Baum, 2003; Boukes and Boomgaarden, 2015; Williams and Delli Carpini, 2011) and more precisely determine political relevance through the eyes of the audience. This also means that future research would have to expand on our explorative findings to come to a more formal description of relevance, in order to provide researchers with a standard that goes beyond relative comparisons.

In conclusion, developing the analysis of second-screen reactions to make inferences about television programs has a lot to offer: While we interpreted the meaning of the most characteristic words manually, it has recently been shown that the topic of tweets about political TV debates can even be inferred automatically using Wikipedia data (Yldrm, *et al.*, 2016). It should thus be feasible to measure the content of television program on a much larger scale. Automated content analysis of textual data is becoming increasingly common in communication science (*e.g.*, Boumans and Trilling, 2016), but the analysis of television content remains difficult and labor intensive. Applying automated content analysis to "community annotation" data from second-screen users seems to be a promising approach to close this gap.

Not only does the current paper have an academic contribution in that it sheds more light on the subtle differences between infotainment programs as being either hard or soft news and how to analyze these, it also carries clear practical implications for media monitoring. This might be particular relevant for public broadcast services (PBS): In democratic states they have, amongst others, the task to prepare citizens for casting informed votes in election (Curran, *et al.*, 2009). This is why PBSs are subsidized by states to inform citizens about politics and current affairs (see, *e.g.*, Soroka, *et al.*, 2012). However, it is very difficult to demonstrate whether programs live up to those requirements or not and deserve subsidies. Whereas now, it is very much in the hands of PBS managers and their (subjective) observations or expensive content analysis to evaluate the public quality of programs, the approach demonstrated in the current paper allows for a faster, more objective, structural and replicable method of assessing the political contribution of media content using the online comments of viewers. Of course, the argument can be extended to the performance of other broadcasters, public or private, on other dimensions (be it political or not): We have shown a simple way to analyze and visualize data that allows conclusions about which aspects of a television program resonate with an audience. 

About the authors

Mark Boukes is a postdoctoral researcher at the Amsterdam School of Communication Research at the University of Amsterdam.
Direct comments to markboukes [at] gmail [dot] com

Damian Trilling is an assistant professor for political communication and journalism in the Department of Communication Science at the University of Amsterdam.
E-mail: d [dot] c [dot] trilling [at] uva [dot] nl

Acknowledgements

This work was carried out on the Dutch national e-infrastructure with the support of the SURF Foundation (<https://www.surf.nl/en>). Both authors contributed equally to the manuscript.

Notes

1. Doughty, *et al.*, 2012, p. 86.

2. Buschow, *et al.*, 2014, p. 141.

3. Jungherr, 2015, p. 181.

4. Interpreting the words to determine relevance requires quite a lot of contextual knowledge. Shamma, *et al.* (2009), for instance, were intrigued by the frequent use of the word “drinking,” which they found out to be present due the fact that many viewers were playing drinking games while watching the debate.

5. This time frame included the Dutch provincial elections on 18 March 2015. However, these were not prominently covered by the media, nor considered very important by the Dutch audience (turnout of 47 percent), and therefore probably has a minimal impact on the findings.

6. Other approaches to determine how characteristic a word is for a given corpus includes the tf-idf (term frequency-inverse document frequency) measure, which makes less sense to use if one has only two corpora to compare. We also experimented with co-occurrence analysis and topic modeling, which did not produce meaningful results due to the shortness of the tweets and the over-representation of few terms.

7. Percentages do not add up to 100 percent because certain words are neither of the coded categories.

8. The studies cited here investigated *Pauw & Witteman*, which is the predecessor of *Pauw* and a very similar talk show program.

References

- Cooper Allen, 2016, "Bernie Sanders really doesn't like Henry Kissinger," *USA Today* (11 February), at <http://www.usatoday.com/story/news/politics/onpolitics/2016/02/11/bernie-sanders-henry-kissinger/80269302/>, accessed 26 March 2017.
- Scott L. Althaus, 2012. "What's good and bad in political communication research? Normative standards for evaluating media and citizen performance," In: Holli A. Semetko and Margaret Scammell (editors). *SAGE handbook of political communication* London: SAGE, pp. 97–112.
- Piet Bakker and Otto Scholten, 2011. *Communicatiekaart van Nederland: Overzicht van media en communicatie*. Eighth edition. Amsterdam: Kluwer.
- Tom Bakker, 2013. "Citizens as political participants: The myth of the active online audience?" Ph.D. thesis, University of Amsterdam, at <http://hdl.handle.net/11245/1.384371>, accessed 26 March 2017.
- Matthew Baum, 2003. *Soft news goes to war: Public opinion and American foreign policy in the new media age*. Princeton, N.J.: Princeton University Press.
- William L. Benoit, Glenn J. Hansen and Rebecca M. Verser, 2003. "A meta-analysis of the effects of viewing U.S. presidential debates," *Communication Monographs*, volume 70, number 4, pp. 335–350.
doi: <http://dx.doi.org/10.1080/0363775032000179133>, accessed 26 March 2017.
- Erik Borra and Bernhard Rieder, 2014. "Programmed method: Developing a toolset for capturing and analyzing tweets," *Aslib Journal of Information Management*, volume 66, number 3, pp. 262–278.
doi: <http://dx.doi.org/10.1108/AJIM-09-2013-0094>, accessed 26 March 2017.
- Linda Bos, Sanne Kruike-meier and Claes de Vreese, 2014. "Gescheiden media? Sociaal-culturele verschillen in mediaconsumptie," In: Mark Bovens, Paul Dekker and Will Tiemeijer (editors). *Gescheiden werelden? Een verkenning van sociaal-culturele tegenstellingen in Nederland*. Den Haag, Netherlands: Sociaal en Cultureel Planbureau, Wetenschappelijke Raad voor het Regeringsbeleid, pp. 235–254.
- Mark Boukes and Hajo G. Boomgaarden, 2015. "Soft news with hard consequences? Introducing a nuanced measure of soft versus hard news exposure and its relationship with political cynicism," *Communication Research*, volume 42, number 5, pp. 701–731.
doi: <http://dx.doi.org/10.1177/0093650214537520>, accessed 26 March 2017.
- Jelle W. Boumans and Damian Trilling, 2016. "Taking stock of the toolkit: An overview of relevant automated content analysis approaches and techniques for digital journalism scholars," *Digital Journalism*, volume 4, number 1, pp. 8–23.
doi: <http://dx.doi.org/10.1080/21670811.2015.1096598>, accessed 26 March 2017.
- Christopher Buschow, Beate Schneider and Simon Ueberheide, 2014. "Tweeting television: Exploring communication activities on Twitter while watching TV," *Communications*, volume 39, number 2, pp. 129–149.
doi: <http://dx.doi.org/10.1515/commun-2014-0009>, accessed 26 March 2017.
- Cédric Courtois and Evelien D'heer, 2012. "Second screen applications and tablet users," *EuroITV '12: Proceedings of the 10th European Conference on Interactive TV and Video*, pp. 153–156.
doi: <http://dx.doi.org/10.1145/2325616.2325646>, accessed 26 March 2017.
- Wouter de Nooy and Jürgen Maier, 2015. "When do attacks work? Moderated effects on voters candidate evaluation in a televised debate," In: Alessandro Nai and Annemarie Walter (editors). *New perspectives on negative campaigning: Why attack politics matters*. Colchester: ECPR Press, pp. 285–304.
- Evelien D'heer and Pieter Verdegem, 2014. "What social media data mean for audience studies: a multidimensional investigation of Twitter use during

a current affairs TV programme," *Information, Communication & Society*, volume 18, number 2, pp. 221–234.
doi: <http://dx.doi.org/10.1080/1369118X.2014.952318>, accessed 26 March 2017.

Nicholas A. Diakopoulos and David A. Shamma, 2010. "Characterizing debate performance via aggregated Twitter sentiment," *CHI '10: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 1,195–1,198.
doi: <http://dx.doi.org/10.1145/1753326.1753504>, accessed 26 March 2017.

Mark Doughty, Duncan Rowland and Shaun Lawson, 2012. "Who is on your sofa?" *EuroITV '12: Proceedings of the 10th European Conference on Interactive TV and Video*, pp. pp. 79–86.
doi: <http://dx.doi.org/10.1145/2325616.2325635>, accessed 26 March 2017.

Myra M. Ferree, William A. Gamson, Jürgen Gerhards and Dieter Rucht, 2002. "Four models of the public sphere in modern democracies," *Theory and Society*, volume 31, number 3, pp. 289–324.
doi: <http://dx.doi.org/10.1023/a:1016284431021>, accessed 26 March 2017.

Julia R. Fox, Glory Koloen and Volkan Sahin, 2007. "No joke: A comparison of substance in *The Daily Show with Jon Stewart* and broadcast network television coverage of the 2004 presidential election campaign," *Journal of Broadcasting & Electronic Media*, volume 51, number 2, pp. 213–227.
doi: <http://dx.doi.org/10.1080/08838150701304621>, accessed 26 March 2017.

Julia R. Fox, James R. Angelini and Christopher Goble, 2005. Hype versus substance in network television coverage of presidential election campaigns, *Journalism & Mass Communication Quarterly*, volume 82, number 1, pp. 97–109.
doi: <http://dx.doi.org/10.1177/107769900508200107>, accessed 26 March 2017.

Daniel Gayo-Avello, 2013. "A meta-analysis of state-of-the-art electoral prediction from Twitter data," *Social Science Computer Review*, volume 31, number 6, pp. 649–679.
doi: <http://dx.doi.org/10.1177/0894439313493979>, accessed 26 March 2017.

Michel M. Haigh and Aaron Heresco, 2010. "Late-night Iraq: Monologue joke content and tone from 2003 to 2007," *Mass Communication and Society*, volume 13, number 2, pp. 157–173.
doi: <http://dx.doi.org/10.1080/15205430903014884>, accessed 26 March 2017.

Andreas Jungherr, 2015. *Analyzing political communication with digital trace data: The role of Twitter messages in social science research*. Cham, Switzerland: Springer.

Andreas Jungherr, Pascal Jürgens and Harald Schoen, 2011. "Why the Pirate Party won the German election of 2009 or the trouble with predictions: A response to Tumasjan, A., Sprenger, T.O., Sander, P.G. and Welpe, I.M. 'Predicting elections with Twitter: What 140 characters reveal about political sentiment'," *Social Science Computer Review*, volume 30, number 2, pp. 229–234.
<http://dx.doi.org/10.1177/0894439311404119>, accessed 26 March 2017.

Andreas Jungherr, Harald Schoen, Oliver Posegga and Pascal Jürgens, 2016. "Digital trace data in the study of public opinion: An indicator of attention toward politics rather than political support," *Social Science Computer Review*, at
<http://journals.sagepub.com/doi/abs/10.1177/0894439316631043>.
<http://dx.doi.org/10.1177/0894439316631043>, accessed 26 March 2017.

Sam N. Lehman-Wilzig and Michal Seletzky, 2010. "Hard news, soft news, 'general' news: The necessity and utility of an intermediate classification," *Journalism*, volume 11, number 1, pp. 37–56.
doi: <http://dx.doi.org/10.1177/1464884909350642>, accessed 26 March 2017.

Jonathan Martin and Patrick Healy, 2015. "In Republican debate, candidates battle sharply on immigration," *New York Times* (10 November), at

<http://www.nytimes.com/2015/11/11/us/republican-debate-fox-business.html>, accessed 26 March 2017.

Jörg Matthes and Adrian Rauchfleisch, 2013. "The Swiss 'Tina Fey effect': The content of late-night political humor and the negative effects of political parody on the evaluation of politicians," *Communication Quarterly*, volume 61, number 5, pp. 596–614.

doi: <http://dx.doi.org/10.1080/01463373.2013.822405>, accessed 26 March 2017.

John McManus, 1995. "A market-based model of news production," *Communication Theory*, volume 5, number 4, pp. 301–338.

doi: <http://dx.doi.org/10.1111/j.1468-2885.1995.tb00113.x>, accessed 26 March 2017.

David T.Z. Mindich, 2005. *Tuned out: Why Americans under 40 don't follow the news*. New York: Oxford University Press.

Newcom Research & Consultancy, 2016. "Nationale social media onderzoek 2016" (24 January), at <http://www.newcom.nl/socialmedia2016>, accessed 26 March 2017.

Danny Oosterveer, 2014. "De laatste cijfers van het socialmediagebruik in Nederland" (13 March), at

<http://www.marketingfacts.nl/berichten/socialmediagebruik-in-nederland-update-maart-2014>, accessed 26 March 2017.

Lukas Otto, Isabella Glogger and Mark Boukes, 2016. "The softening of journalistic political communication — A critical review of sensationalism, soft news, infotainment, and tabloidization and a comprehensive framework model," *Communication Theory*, at

<http://onlinelibrary.wiley.com/doi/10.1111/comt.12102/abstract>, accessed 24 March 2017.

doi: <http://dx.doi.org/10.1111/comt.12102>, accessed 26 March 2017.

Matthew Pittman and Alex C. Tefertiller, 2015. "With or without you: Connected viewing and co-viewing Twitter activity for traditional appointment and asynchronous broadcast television models," *First Monday*, volume 20, number 7, at <http://firstmonday.org/article/view/5935/4663>, accessed 26 March 2017.

doi: <http://dx.doi.org/10.5210/fm.v20i7.5935>, accessed 26 March 2017.

Racine Group, 2002. "White paper on televised political campaign debates," *Argumentation and Advocacy*, volume 38, number 4, pp. 199–218.

Paul Rayson and Roger Garside, 2000. "Comparing corpora using frequency profiling," *WCC '00: Proceedings of the Workshop on Comparing Corpora*, volume 9, pp. 1–6; version at

http://ucrel.lancs.ac.uk/people/paul/publications/rg_acl2000.pdf, accessed 24 March 2017.

doi: <http://dx.doi.org/10.3115/1117729.1117730>, accessed 26 March 2017.

Carsten Reinemann, James Stanyer, Sebastian Scherr and Guido Legnante, 2012. "Hard and soft news: A review of concepts, operationalizations and key findings," *Journalism*, volume 13, number 2, pp. 221–239.

doi: <http://dx.doi.org/10.1177/1464884911427803>, accessed 26 March 2017.

Reuters Institute for the Study of Journalism. 2016. "Reuters Institute digital news report 2016," at

<http://reutersinstitute.politics.ox.ac.uk/sites/default/files/Digital-News-Report-2016.pdf>, accessed 26 March 2017.

Jason Rittenberg, David Tewksbury and Shanna Casey, 2012. "Media preferences and democracy: Refining the 'relative entertainment preference' hypothesis," *Mass Communication and Society*, volume 15, number 6, pp. 921–942.

doi: <http://dx.doi.org/10.1080/15205436.2011.622065>, accessed 26 March 2017.

Franziska S. Roth, Carina Weinmann, Frank M. Schneider, Frederic R. Hopp and Peter Vorderer, 2014. "Seriously entertained: Antecedents and consequences of hedonic and eudaimonic entertainment experiences with political talk shows on TV," *Mass Communication and Society*, volume 17, number 3, pp. 379–399.

doi: <http://dx.doi.org/10.1080/15205436.2014.891135>, accessed 26 March

2017.

Philip Rucker, 2015. "Democrats see a more substantive, if sleepy, debate than rowdy GOP show," *Washington Post* (12 October), at https://www.washingtonpost.com/politics/democrats-see-a-more-substantive-if-sleepy-debate-than-rowdy-gop-show/2015/10/12/f1b114da-7070-11e5-9cbb-790369643cf9_story.html, accessed 26 March 2017.

David A. Shamma, Lyndon Kennedy and Elisabeth F. Churchill, 2010. "Tweetgeist: Can the Twitter timeline reveal the structure of broadcast events?" *CSCW Horizons*, pp. 589–593; version at <https://research.yahoo.com/publications/5852/tweetgeist-can-twitter-timeline-reveal-structure-broadcast-events>, accessed 26 March 2017.

David A. Shamma, Lyndon Kennedy, and Elizabeth F. Churchill, 2009. "Tweet the debates: Understanding community annotation of uncollected sources," *WSM '09: Proceedings of the First SIGMM Workshop on Social Media*, pp. 3–10. doi: <http://dx.doi.org/10.1145/1631144.1631148>, accessed 26 March 2017.

Sean Sullivan, 2016. "Republican debate: Raised voices, name calling and personal attacks," *Washington Post* (13 February), at https://www.washingtonpost.com/politics/all-eyes-will-be-on-trump-at-republican-debate/2016/02/13/bd30bae2-d274-11e5-abc9-ea152f0b9561_story.html, accessed 26 March 2017.

Damian Trilling, 2015. "Two different debates? Investigating the relationship between a political debate on TV and simultaneous comments on Twitter," *Social Science Computer Review*, volume 33, number 3, pp. 259–276. doi: <http://dx.doi.org/10.1177/0894439314537886>, accessed 26 March 2017.

Damian Trilling and Klaus Schoenbach, 2015. "Investigating people's news diets: How online news users use offline news," *Communications: European Journal of Communication Research*, volume 4, number 1, pp. 67–91. doi: <https://doi.org/10.1515/commun-2014-0028>, accessed 26 March 2017.

Maurice Vergeer and Philip H. Franses, 2016. "Live audience responses to live televised election debates: Time series analysis of issue salience and party salience on audience behavior," *Information, Communication & Society*, volume 19, number 10, pp. 1,390–1,410. doi: <http://dx.doi.org/10.1080/1369118X.2015.1093526>, accessed 26 March 2017.

Paul Hendriks Vettehen, Johannes Beentjes, Koos Nuijten and Allerd Peeters, 2011. "Arousing news characteristics in Dutch television news 19902004: An exploration of competitive strategies," *Mass Communication and Society*, volume 14, number 1, pp. 93–112. doi: <http://dx.doi.org/10.1080/15205431003615893>, accessed 26 March 2017.

Bruce A. Williams and Michael X. Delli Carpini, 2011. *After broadcast news: Media regimes, democracy, and the new information environment*. New York: Cambridge University Press.

Anke Wonneberger, Klaus Schoenbach and Lex van Meurs, 2011. "Interest in news and politics — or situational determinants? Why people watch the news," *Journal of Broadcasting & Electronic Media*, volume 55, number 3, pp. 325–343. doi: <http://dx.doi.org/10.1080/08838151.2011.597466>, accessed 26 March 2017.

Ahmet Yldrm, Suzan Üsküdarlı and Arzucan Özgür, 2016. "Identifying topics in microblogs using Wikipedia," *PLoS ONE*, volume 11, number 3, e0151885, at <http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0151885>, accessed 26 March 2017. doi: <http://dx.doi.org/10.1371/journal.pone.0151885>, accessed 26 March 2017.

Barbie Zelizer, 2007. "Introduction: On finding new ways of thinking about journalism," *Political Communication*, volume 24, number 2, pp. 111–114. doi: <http://dx.doi.org/10.1080/10584600701312829>, accessed 26 March 2017.

Editorial history

Received 6 October 2016; revised 8 January 2017; accepted 26 March 2017.



“Political relevance in the eye of the beholder: Determining the substantiveness of TV shows and political debates with Twitter data” by Mark Boukes and Damian Trilling is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International License](https://creativecommons.org/licenses/by-sa/4.0/).

Political relevance in the eye of the beholder: Determining the substantiveness of TV shows and political debates with Twitter data by Mark Boukes and Damian Trilling.

First Monday, Volume 22, Number 4 - 3 April 2017

<http://journals.uic.edu/ojs/index.php/fm/article/view/7031/6093>

doi: <http://dx.doi.org/10.5210/fm.v22i4.7031>



A Great Cities Initiative of the University of Illinois at Chicago [University Library](https://www.library.uic.edu/).

© *First Monday*, 1995-2017. ISSN 1396-0466.