Through the eyes of others

Social Opinion Monitoring in a televised debate

Working Paper

prepared for the Joint Annual Conference 2019
of

AK „Politik und Kommunikation“ (DVPW),
FG „Kommunikation und Politik“ (DGPuK) und
FG „Politische Kommunikation“ (SGKM)

February 7-9, 2019

Simon Richter, M.A.

**Freie Universität Berlin
Otto-Suhr-Institut**

simon.richter@fu-berlin.de

Disclaimer: This is work in progress. Please do not redistribute without permission.

Abstract:

More and more people engage in social media platforms while watching televised debates during election campaigns. This study explores the potential of biases in social media users' comments in influencing the debate viewing experience. We argue people do rely on biases in these additional information environments, in order to make sense out of a televised debate. To empirically assess these effects, the paper builds on data from a survey experiment on a German televised debate. While evidence for effects on the perceived debate performance of the candidates are rather scarce, we find consistent evidence for people drawing on a bias as an indicator for general public's evaluation of the candidates' debate performance. Furthermore, the paper highlights the differential roles positive and negative biases play for the above-mentioned effects.

The televised debate in the run-up to German national election in 2017 was accompanied by a sizeable number of commentaries on the social media network Twitter (M. König & W. König, 2017). While debates in previous debates as well induced simultaneous social media discourses, this time something was different: A group of online activists tried to hijack the Twitter discourse. As uncovered by a BuzzFeed journalist (Schmehl, 2017), these activists planned their attack in anonymous online forums beforehand. They aimed to establish an alternative hashtag (#verräterduell) and posted aggressive messages, which should deter voters from the two debating candidates, chancellor Angela Merkel from the Christian Democratic Union and her main challenger Martin Schulz (Social Democratic Party). Although the group, which became known as “Reconquista Germanica” (Von Hammerstein, Höfner, & Rosenbach, 2017), was not able to overtake the social media discourse during the debate, the hashtag was clearly visible (Grimme, Assenmacher, Preuss, Adam, & Lütke-Stokdiek, 2017) and their messages potentially changed the tone of Twitter discussion which accompanied the televised debate.

This incident strikingly illustrates the fact that we do have to rethink political communication. How people inform themselves about politics has changed in the past decades. The undisputed monopoly of the traditional mass media in distributing political information is getting challenged by what Chadwick and colleagues have labeled the “hybrid media system” (2016). The emergence and establishment of digital media led to new forms of communications in which the logics of the traditional mass media and digital media interact. This change potentially entails plenty of consequences for how citizens nowadays form their views of the world that surrounds them in general (Neuman, 2016), and how they inform themselves about politics in particular.

This also holds for mediated communication during electoral campaigns. How people watch televised debates, which are the prominent central media events during electoral campaigns (McKinney & Spialek, 2017), is affected by this new hybridity of the media system. The ubiquity of mobile devices and the increasing levels of internet access and usage led more and more people to simultaneously engage in following – or even contributing – to debate-related discourses on social media channels while watching the televised debate (Chadwick, 2017; Gil de Zúñiga & Liu, 2017; Gottfried, Hardy, Holbert, Winneg, & Jamieson, 2017; Vaccari, Chadwick, & O'Loughlin, 2015).

Recent studies provide evidence for the claim that a social media discourse may affect resulting debate attitudes (Hahn, H.-Y. Lee, Ha, S. Jang, & J. Lee, 2017; Houston, Hawthorne, Spialek, Greenwood, & McKinney, 2017; McKinney, Houston, & Hawthorne, 2014). However, we do not yet understand how these effects originate. This is exemplified in the rather inconsistent effect patterns we find – even within single studies. For instance, Hahn and colleagues do find evidence for slanted tweets to induce bandwagon as well as underdog effects (Hahn et al., 2017, p. 14).

This calls for a more detailed look into how and under which circumstances a bias in this type of additional information environment is actually able to influence the debate viewing experience. Therefore, this paper proposes to widen the focus of the existing research in two ways. On the one hand, we think one should also include a third person perspective on the dependent variable side. We argue that a bias in concurrently consumed social media messages might influence in the first place what debate watchers think the public opinion about the debate and its contenders might look like. On the other hand, we will focus on the specific nature of the bias in these commentaries and what consequences these entail. We argue that negative and positive biases might operate quite distinctively, when it comes to the formation of post-debate attitudes, e.g. who performed best in the debate.

In the remainder of the paper, we give a short literature review and introduce our extended theoretical model. We then formulate some hypotheses that are tested drawing on a survey experiment, executed in the weeks before the German national election in 2017. A conclusion section will infer some learnings for the further research.

Literature Review

Doubtless, televised debates are among the most popular single media events during an electoral campaign. No other event during a campaign attracts a comparable amount of media attention and audiences of similar size. In the US, the first Clinton-Trump debate in the 2016 general election campaign reached roughly 84 million watchers – a new record figure in debate audience size (McKinney & Spialek, 2017). This also holds true for German debates. As in previous electoral campaigns (Faas, J. Maier, & M. Maier, 2017), the debate in the run-up to the German national election in 2017 was the single most watched election broadcast with an estimated audience of 16.3 million people (Gscheidle, Geese, & Gerhard, 2017).

To date scholars from a variety of different disciplines accumulated an impressive amount of research on televised debates and their reception (McKinney & Spialek, 2017). One set of studies focuses on how informational contexts do influence the perception of a debate. These start from the observation that televised debates as such are not isolated events. Although they are outstanding compared to other media events during electoral campaigns, they are still embedded in informational contexts, which may influence how viewers perceive the debate. These contexts include the extensive mass media coverage before and after the debate (G. Lang & K. Lang, 1979; Tsfati, 2003), as well as interpersonal communication in the aftermath (J. Maier & Faas, 2003; Tsfati, 2003) or with fellow co-watchers during the debate (Leuchte, 2017).

While the existence of such effects has been broadly discussed, the underlying causal mechanism is not well understood. Some argue that televised debates are by nature complex events as they involve a range of extensive policy discussions and a high degree of conflict among the contenders. This might create a need for additional information facilitating orientation and guidance to make sense of the debate (Chadwick, O’Loughlin, & Vaccari, 2017; Tsfati, 2003). In fact, Fein and colleagues demonstrated in a series of experiments that even some rough information about others’ opinions – for example cheers of unknown fellow co-watchers during a debate – may influence one’s own judgements about candidates’ performances (Fein, Goethals, & Kugler, 2007). In another study Davis and colleagues demonstrated how the graphical representation of the net opinion on the debaters’ performance of unknown others might influence the debate viewers’ judgement about who won the debate (Davis, Bowers, & Memon, 2011).

The media landscape has changed drastically within the past decade. The emergence of new forms of communication technology – first and foremost the omnipresence of the internet – changed how the media system is composed and works nowadays. Chadwick and colleagues have labeled this new, changed media system as “hybrid” (Chadwick et al., 2016), hinting on the intertwined media logics of the traditional mass media and the novel digital media.

The most tangible phenomenon emerging from this new hybrid media system might be the practice of second screening. This refers to a “bundle of practices that involve integrating, and switching across and between, broadcast media and social media” (Vaccari et al., 2015, p. 1044) and has as elsewhere been described as “social watching” (McKinney et al., 2014; Thorson, Hawthorne, Swasy, & McKinney, 2015) or “dual screening” (Vaccari et al., 2015). It basically

complements elements of the traditional mass media (i.e. a television broadcast) with a new communication channel from the world of digital media.

As one can deduce from the broad definition of second screening it may involve quite different forms of interaction with social media channels. Vaccari et al. describe two styles of second screening practices: those being “lean-forward”, which involves “relatively active, purposive information-seeking and information-production practices” (Vaccari et al., 2015, p. 1044) and a “lean-backward” way of second screening, where people engage in “relatively passive, information-reception practices classically associated with broadcast media” (ibid.). Within the scope of this paper we understand second screening as concurrent consumption of social media content related to an audiovisual content, thereby ignoring behavior that involves active contribution to the social media discourse.

Second screening is most popular during media events (Anstead & O’Loughlin, 2011; Y.-R. Lin, Keegan, Margolin, & Lazer, 2014; Vaccari et al., 2015) such as televised debates (Jungherr, 2015; Y.-R. Lin et al., 2014; Trilling, 2014). During the first debate between Barack Obama and Mitt Romney in 2012 an estimated 10.3 million tweets about the debate have been sent (McKinney et al., 2014). The estimates are lower but still of considerable size for German debates with estimates of 160 000 Tweets (Nyhuis & Friedrich, 2017) for the 2013 debate between Angela Merkel and Peer Steinbrück and 194 000 Tweets for the most recent 2017 debate between Angela Merkel and Martin Schulz (M. König & W. König, 2017).

Research focused on the content of social media discourses during televised debates yielded consonant evidence indicating the close connection between the social media discourse and the actual content of the broadcast in terms of addressed issues (Anstead & O’Loughlin, 2011; Nyhuis & Friedrich, 2017; Shah et al., 2015; Trilling, 2014). Furthermore, these studies hint on commentaries being rather negative (Diakopoulos & Shamma, 2010; Nyhuis & Friedrich, 2017) and marked by humor or irony (Freelon & Karpf, 2014; Trilling, 2014).

Another strand of research focuses on the second screeners themselves. Gil de Zúñiga and Liu (2017) show that younger people tend to second screen for political content more than older people. They do so in order to get further information on and discuss about the broadcast content (Chadwick et al., 2017; Gil de Zúñiga, Garcia-Perdomo, & McGregor, 2015). In the case of televised debates the need to gauge the opinion of the public has been shown to be a recurrent motive for second screen usage (Gorkovenko & Taylor, 2016).

Finally, a growing body of research on second screening in political contexts looks at attitudinal and behavioral consequences. It has been shown that second screening political content is linked to higher levels of political participation (Gil de Zúñiga et al., 2015; T. Lin & Chiang, 2017), issue learning (Houston et al., 2017) and self-reported openness to political persuasion (Barnidge, Gil de Zúñiga, & Diehl, 2017). Of particular interest for the present paper are the few publications on the influence of second screening on the perception of televised debates. To our knowledge this is limited to a paper by McKinney et al. (2014) and a paper by Hahn et al. (2017), which both give some evidence that an additional informational context in form of other social media users' opinions may actually influence the evaluation of the candidates' performance (Hahn et al., 2017; McKinney et al., 2014) and – to some degree – even the vote choice (McKinney et al., 2014).

However, we do not have a solid understanding of how and under which circumstances these additional information environments are able to influence the debate viewing experience. Is it indeed the slant of the second screen that affects how viewers perceive the debate? The study of Hahn and colleagues thus far presents the most solid data to look for such effects of biases in these second screen environments on the evaluation of the debaters' performances (Hahn et al., 2017). Using real-time response measurements in an experimental setting, they are able to link the temporal changes of a bias in a social media discourse they present to their subjects to their real-time feedback on the contenders' performances during the debate. However, their experimental design, consisting in one control group watching the debate without a second screen and an experimental group provided with a sample of tweets, does actually not allow them to attribute the over time variation of the slanted second screen to the evaluation of the debaters' performance. Technically they do compare those only watching a televised debate with those who are provided with an additional information environments. We argue a third experimental group would have been necessary to assure that it is actually the slant of the social media messages rather than the pure provision of the additional information environment that drives the performance evaluations. This third experimental group should have provided a second differentially biased stream of social media messages.

The differential effects from biased second screens

This paper departs from these observations and aims to assess whether and under which conditions a bias in a set of social media messages consumed while watching a televised debate is able to influence post-reception attitudes. Therefore we draw on the social influence

literature, which commonly is traced back to Sherif’s autokinetic effect experiment (Sherif, 1935), the Asch conformity experiments and its variations (Asch, 1951, 1955; Deutsch & Gerard, 1955). The original experiments demonstrated that subjects show a tendency to conform with stances expressed by a majority group. It is theorized that this conformity tendency is due to a normative motive (to avoid isolation) or due to informational reasons (because the majority opinion is assumed to be correct) (Deutsch & Gerard, 1955).

While these classical experiments investigated social conformity effects on non-political issues, Levitan and Verhulst recently brought up evidence for existing conformity effects on political attitudes. In two experiments they showed that the simple knowledge about others political viewpoints might actually lead to attitude conversion (Levitan & Verhulst, 2016). In a series of experiments on televised debates, Fein and colleagues demonstrated that judgements of candidate performances could be influenced through reactions of the audience, like cheers, laughter or visually fabricated feedback of other audience members (Fein et al., 2007). They close their paper with the remark that the communication via internet might play an important role in future televised debates, as “very large numbers of citizens, particularly younger voters, will provide spontaneous audience reactions, [...] by posting to blogs, electronic forums and boards, and so on” (Fein et al., 2007, p. 189). In this sense, statements on a televised debate transmitted through social media channels might very well alter the way debate viewers judge the candidates’ performances. We consequently formulate the hypothesis:

H1a: A bias in the valence of the concurrently consumed social media messages shapes a subject’s personal evaluation of the candidates’ debate performances (in the direction of the bias).

However, the social media comments of other debate viewers might not only affect subjects’ personal evaluation of the candidates’ performances, but also the perceived public opinion on the debaters’ performance. This is the general idea that the comments on co-watchers might be used as an information shortcut on whether the candidates have performed well in the eyes of some “generalized others” (Mutz, 1998).

Although one could argue that some individual comments of social media users are quite unlikely to be taken as an indicator for public opinion, research advises us otherwise. In fact, statements of individuals might actually be a more powerful information cue for public opinion than ‘base rate’ information (e.g. polls). Exemplification theory suggests so-called exemplars

(in our case fellow co-watchers tweeting during a televised debate) to be quite influential when it comes to social judgement as e.g. gauging the public opinion on an issue or a particular candidate (Zillmann, 2002). Although exemplars seldom reflect the actual opinion distribution among a given universe of persons, they “retain their psychological superiority due to their vividness and memorability” (Waddell & Bailey, 2017, p. 3) as compared to the bare numbers of an anonymous poll. Daschmann shows that exemplars used by journalists to illustrate public opinion (labeled vox pops) are more effective in influencing the perceived opinion climate than poll results (2000). In fact, vox pops were even able to influence personal opinions whereas poll results were not. At the same time, messages from Twitter are often used by journalists as vox pops in order to illustrate public opinion (Beckers & Harder, 2016).

Also user-generated comments from the internet that are not curated in a mass media frame and presented by journalists as public opinion examples are able to influence the perception of public opinion. For instance Lee and Jang (2010) found user comments on an online news article were able to influence readers’ perception of the general opinion climate on an certain issue. These finding are further corroborated by Zerback and Fawzi, who show that a “relatively low number of online exemplars considerably influence perceived public support for the eviction of violent immigrants” (2017, p. 1034). We therefore expect the bias of social media messages evaluating the performance of a televised debate’s candidates to influence the estimated public opinion on the debaters’ performances. Stated as an hypothesis:

H1b: A bias in the valence of the concurrently consumed social media messages shapes a subject’s perception of the general public’s evaluation of the candidates’ debate performance (in the direction of the bias).

In his seminal book on communication effects research in the digital age, W. Russell Neuman highlights the importance of communication valence as a subject of investigation. He claims that all “human communication, especially in the public sphere, tends to be valanced” (Neuman, 2016, p. 46) and scholars should therefor pay more attention to this valence. In fact, we do observe a prominent share of valanced messages in the social media discussions during media events like televised debates. For instance, Diakopoulos and Shamma identified around two thirds of the tweets during a televised debate between John McCain and Barack Obama in 2008 to be either distinctly negative or positive in its tone (2010). The vast majority (more than 40 per cent of all tweets) were identified as negative messages. This coincides with the

analysis of Nyhuis and Friedrich, who detected a clear negative bias in the tweets on Angela Merkel during the German televised debate in 2013 (Nyhuis & Friedrich, 2017).

Given these observations, this paper sets a special focus on the specific valence of social media users' comments and its consequences on information processing and attitude formation of a second screening debate viewer. To put it into a simple question, we may ask whether negative comments are as influential as equally positive comments. Especially when talking about second screening, a form of media consumption, which involves a high degree of multitasking (van Cauwenberge, Schaap, & van Roy, 2014), we should assume that some messages might be more successful in attracting the viewers' attention and bind more processing effort than others. The negativity effect (Ito, Larsen, N. Smith, & Cacioppo, 1998; Rozin & Royzman, 2001) describes negative messages as being more successful in attracting attention, being recalled and even being persuasive as compared to positive information. Along these lines, an eye-tracking study simulating a second-screening situation while watching TV news detected subjects to focus significantly longer on negative comments on the second screen than they focused on positive comments (Kätsyri, Kinnunen, Kusumoto, Oittinen, & Ravaja, 2016). This leads us to hypothesize:

H2: The expected effects of social media messages on subjects' attitudes (stated in H1a and H1b) are stronger when social media messages are biased negatively rather than positively.

Data & Methods

In the run up to the German federal elections in 2017 the sitting chancellor Angela Merkel confronted her challenger Martin Schulz in a 90-minutes televised debate three weeks before election day. More than 16 million viewers watched the debate (Gscheidle et al., 2017). It was not only accompanied by a vast amount of mass media coverage before and after the event, but also by a vivid discussion on social media platforms – especially while the debate took place.

For the present study we asked people to assess a three and half minutes long extract of the debate which we introduced as one of the key sequences of the debate. In the selected scene, the two contenders talk about how the German government should position itself towards Turkey. Both Angela Merkel (1:56 Minutes) and Martin Schulz (1:31 Minutes) speak for about the same period of time. Additionally, we showed a series of 12 tweets in the bottom section of

the video to our subjects (see Figure 1). These tweets were original tweets sent during the debate and were selected to match with the scene in terms of content and timing. The subjects were not informed about the tweets beforehand in order not to artificially highlight the importance of the tweets for the study.

Figure 1: Extract from a video stimulus

The tweets shown to the subjects were systematically varied between five groups. From a bigger dataset of tweets captured during the televised debate, tweets were selected that showed a clear valence towards one of the two candidates being either pro or contra the respective candidate. The tweets were displayed to the subjects following the patterns outlined in Table 1. This resulted in a net neutral set of tweets, which will serve as our unbiased control group, and four biased sets of tweets. The biased sets focused on one candidate by devoting eight out of 12 tweets to one candidate. Of these eight tweets six either evaluated the candidate positively or negatively, the remaining two tweets evaluated the candidate vice versa in order to conserve credibility and warrant external validity. The remaining four tweets concerned the other respective candidate and consistent in two positive and two negative messages towards that other candidate.

Please note that all tweets used in this study were authentic tweets sent during the debate. We secured to have the largest possible share of common tweets among the experimental groups. In fact, 8 of 12 tweets were identical in all five conditions. This secures the comparability of our stimuli. With the remaining 4 tweets, we had enough variability to induce perspicuous biases. Each tweet was displayed for 15 seconds. They were formatted to closely resemble the original look of tweets shown on Twitter.

Table 1: Composition of stimuli

Groups	Tweets			
	Pro-Merkel	Contra-Merkel	Pro-Schulz	Contra-Schulz
Neutral	3	3	3	3
Merkel Positive	6	2	2	2
Merkel Negative	2	6	2	2
Schulz Positive	2	2	6	2
Schulz Negative	2	2	2	6

The experimental study followed a between-subject design, thus the survey participants were randomly assigned to one of the five video versions. Directly after watching the video, subjects were asked to evaluate the candidates' performances¹ (in terms of persuasiveness) and to state how they believe the general debate audience evaluated the candidates' performances² (in terms of persuasiveness) in the specific scene. Subsequently the survey asked the subjects to assess whether the tweets were either positive or negative towards the candidates, which allows us to check for successful manipulation³. Furthermore, subjects were asked for their party identification, age, gender and formal degree of education.

The study was part of a bigger online survey fielding in the two weeks before the election. The participants were recruited via social media advertisements, a press release, newsletters and mailing lists. As an incentive for participation, those willing to give their email took part in a raffle of amazon vouchers. The resulting convenience sample comprised 1106 completed surveys. The survey software allowed to check whether the subjects watched the complete video. In order to control for stimulus compliance, we sorted out those who did not watch the whole video. Additionally, we exclude participants who lasted longer than 10 minutes in watching the video. This left us with data from 699 participants.

¹ The question asked the subject to separately evaluate whether the two candidate „convinced“ them in the scene: “Einmal ganz allgemein gesprochen, wie überzeugend fanden Sie Angela Merkel bzw. Martin Schulz in dem eben gesehenen Ausschnitt?” Antwortskala: 1 „Überhaupt nicht überzeugend“ 5 „Sehr überzeugend“.

² Exact question wording: „Was glauben Sie, wie überzeugend fanden die Zuschauerinnen und Zuschauer des TV-Duells insgesamt Angela Merkel bzw. Martin Schulz in dem eben gesehenen Ausschnitt?“ Antwortskala: 1 „Überhaupt nicht überzeugend“ 5 „Sehr überzeugend“.

³ Exact question wording: „In dem Video, dass Sie sich angesehen haben, wurden einige Beiträge von Twitter, sogenannte Tweets, eingeblendet. Wenn Sie an diese Tweets denken: Wie wurden Angela Merkel bzw. Martin Schulz Ihrer Meinung nach alles in allem in diesen Beiträgen dargestellt?“ Antwortskala: 1 „Sehr negativ“ 5 „Sehr positiv“.

Although based on a convenience sample, the pool of subjects participating in our study was more diverse than an ordinary student sample. The age of our subjects ranged from 18 to 80 years, with a median age of 28 years. However, our sample was rather highly educated with around 90 percent of the participants having completed high school. In terms of gender we had about 44 percent women participating in the survey, while 56 percent were men. Turning to partisanship, we registered more people feeling connected to the SPD than to the CDU/CSU (22 percent of the total sample versus 14 percent) while the vast majority of our sample did identify with another party or reported to be independent (64 percent).

Table 2: Composition of the sample

Variable	Groups					Total
	Neutral	Merkel Positive	Merkel Negative	Schulz Positive	Schulz Negative	
n	145	137	135	126	156	699
Age	35.3 (15.4)	36.0 (14.3)	36.1 (15.6)	35.1 (15.3)	35.6 (13.8)	35.6 (14.8)
Female	48%	49%	43%	43%	38%	44%
Formal education (high school)	88%	90%	87%	89%	90%	89%
Party identification						
CDU/CSU	16%	14%	21%	10%	11%	14%
SPD	24%	23%	15%	25%	20%	22%
None or other	60%	63%	64%	65%	69%	64%

A look on Table 2 reveals that randomization worked well. In the critical case of the distribution of party identification among our experimental groups, we double checked for possible dependencies. Although a chi-square test did not bring forward any evidence for a systematic bias between the groups ($X^2 = 11.919$, $df = 8$, $p\text{-value} = 0.1548$), all models estimated in the remainder of the paper will control for party identification.

Results

Before we test our hypothesis, we use OLS regressions to check whether our stimulus manipulation was successful (Table 3). We analyze the perceived tone of the Twitter messages (measured on a five-point scale) for each candidate separately. The “Neutral” group serves as our control group and consequently is the reference category for our effect estimates.

As the estimates indicate the stimulus manipulation was perceived as biased in the intended direction. The Twitter tone towards Angela Merkel was perceived to be significantly more positive in the “Merkel Positive” group and significantly more negative in the “Merkel Negative” group (as compared to the “Neutral” group respectively). Please note that there are no significant differences in perceived Twitter tone towards Merkel between the two Schulz groups and the “Neutral” group as Pro- and Contra-Merkel tweets were balanced in all three groups.

Table 3: Perceived Twitter Tone by Candidates (Stimulus Check)

Results		
	Perceived Twitter Tone	
	Merkel	Schulz
Merkel Positive	0.576*** (0.107)	-0.131 (0.108)
Merkel Negative	-0.464*** (0.109)	0.323** (0.109)
Schulz Positive	0.072 (0.114)	0.735*** (0.112)
Schulz Negative	0.057 (0.110)	-0.633*** (0.109)
CDU/CSU	0.029 (0.102)	-0.123 (0.103)
SPD	0.070 (0.083)	-0.053 (0.083)
Constant	3.423*** (0.084)	3.097*** (0.083)
N	632	635
R-squared	0.136	0.212
Adj. R-squared	0.127	0.205
OLS-Regression; * p<0.05; ** p<0.01; *** p<0.001		

For Martin Schulz we do observe the identical, expected patterns. However, there is an additional significant difference in the perceived Twitter tone between the “Merkel Negative” and the “Neutral” condition. Although subjects in both groups did receive a balanced set of

negative and positive tweets concerning Schulz, the negative bias towards Merkel seems to have let the subjects to perceive a relatively more friendly tone towards Schulz in the tweets.⁴

Turning to the personal evaluation of the candidates' performances (Table 4), we do only find one significant effect estimate for the experimental stimuli: in the case of Angela Merkel's debate performance those in the "Merkel Positive" condition were more convinced by Angela Merkel than those in the "Neutral" group. Although the (negative) effect estimate for the "Merkel Negative" group does clearly point into the expected direction, it does not reach conventional levels of significance. In the case of the perceived performance of Martin Schulz, we do not find any significant differences between the "Neutral" and the biased Schulz conditions. Hypothesis 1a therefore has to be rejected.

Yet, re-estimating the models with the respective negative bias groups (e.g. "Merkel Negative" and "Schulz Negative") as the reference groups, we do find significant groups differences in the expected direction (see Table 5 in the Appendix). This is to say, comparing the two biased groups, we do find solid evidence for some sort of conformity effects among our subjects.

It does not come as a surprise that ideological predispositions show clear effects on the evaluation of the candidates' performances. Those identifying with the Merkel's Christian Democratic Union clearly evaluate her as performing better than other respondents. And those who feel attached to social democrats evaluate Schulz as performing better than other subjects.

⁴ Interestingly, we do not observe any projection effects. Identifying with one of contenders' parties did not lead to biased perception of the Twitter tone.

Table 4: Explaining performance evaluation and Perceived public opinion towards the candidates

Results				
	Performance evaluation		Perceived opinion of others	
	Merkel	Schulz	Merkel	Schulz
Merkel Positive	0.286*	-0.075	0.320**	-0.101
	(0.129)	(0.125)	(0.106)	(0.106)
Merkel Negative	-0.141	-0.007	-0.094	0.121
	(0.130)	(0.125)	(0.106)	(0.106)
Schulz Positive	0.044	0.166	0.085	0.331**
	(0.132)	(0.127)	(0.108)	(0.108)
Schulz Negative	0.040	-0.084	0.198	-0.025
	(0.126)	(0.121)	(0.102)	(0.102)
CDU/CSU	0.686***	-0.184	0.227*	0.030
	(0.120)	(0.116)	(0.098)	(0.098)
SPD	-0.208*	0.850***	0.075	0.195*
	(0.102)	(0.098)	(0.084)	(0.084)
Constant	3.587***	3.135***	3.543***	3.060***
	(0.097)	(0.093)	(0.079)	(0.079)
N	697	697	692	692
R-squared	0.070	0.120	0.033	0.035
Adj. R-squared	0.062	0.112	0.024	0.026

OLS-Regression; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

A much clearer, because candidate-invariant effect pattern emerges once we focus on the perceived opinion of others. Those being provided with the positively biased set of tweets of the respective candidate did assume that the public opinion about the candidate's performance is more positive than in the "Neutral" group. There is no such effect observable for those in the negatively biased conditions. This provides partial evidence for Hypothesis 1b. The significant differences in the party identification dummies point towards projection effects. Feeling connected to a candidates' party made the subjects more optimistic about the public opinion on the candidates' performance.

The negativity hypothesis (H2) expected the negative bias groups ("Merkel Negative" and "Schulz Negative") to yield stronger effects than the positive bias groups ("Merkel Positive" and "Merkel Negative") on both the evaluation of the candidates' performances by the subjects

themselves and the perceived public opinion on these evaluations. In fact, we only find the positively biased conditions (as compared to the “Neutral” condition) to have an effect on the post-reception attitudes. The negatively biased conditions did not lead to systematically different post-reception attitudes than in the “Neutral” group. Hypothesis 2 therefore must be rejected.

Discussion

We then can conclude that debate viewers are sensitive to biases in social media commentaries read while watching a debate. This ultimately gets corroborated when we compare the outcomes not only between the bias groups and the neutral, but also between the respective bias groups within the candidates (e.g. “Schulz Negative” vs. “Schulz Positive”, see Table 5 in the Appendix). However, we started this paper by asking whether a biased (as compared to an unbiased) set of social media messages consumed while watching an extract of a televised debate is actually able to influence post-debate evaluations. Sticking to this theoretical approach, we do only find sparse effects of biases on the debate viewers’ perception of the candidate performances (H1a). Effect patterns do get clearer (and slightly more pronounced) when we turn to the assessment of the opinion climate of the candidate performances (H1b). This might point towards underlying information processes these commentaries are evoking in the viewers’ heads. In line with the exemplification theory, the voices of some unknown fellow co-watchers seem to be able to inform the subjects’ beliefs about how the general public did evaluate the candidates’ appearance. As a consequence of this finding, further research should have a more detailed look at the concept of “the others” to whom the perceived opinion climate is referenced. Would changing the reference group (opinion of other debate watchers versus social media users versus own peers) yield different effects? Answering this question lays beyond the scope of the present paper and should be subject to further investigation.

Not all biases are alike. The results of our experiment impressively demonstrate the importance to take the valence of a message into account when doing communication effect research. The effect patterns across candidates and dependent variables suggest differential potentials of negative and positive biases to influence debate attitudes. However, it would be premature to conclude that only positive biases in other users commentaries do change debate attitudes. This is because the “missing” effects of negative stimuli might actually be rooted in the nature of our “Neutral” condition as control group. This set of tweets is not composed of actual neutral tweets, but by balanced amounts of negative and positive comments regarding each candidate.

Assuming differential perception and processing patterns for positive and negative messages, we might actually observe what Rozin and Royzman labeled a “negativity dominance” effect. Following the authors, “the holistic perception and appraisal of integrated negative and positive events (or objects, individuals, hedonic episodes, personality traits, etc.) is more negative than the algebraic sum of the subjective values of those individual entities” (Rozin & Royzman, 2001, 298f).

If this holds true for the arithmetically balanced “Neutral” condition, its tweets should have been perceived as being slightly more negative in the tone towards the two candidates. In fact, our subjects perceived Martin Schulz to be evaluated perfectly neutral (mean = 3.05, sd = 0.93) in the Tweets of the “Neutral” group. Angela Merkel even is perceived as being evaluated slightly positively in the messages of the “Neutral” group (mean = 3.43, sd = 0.91).⁵ While these figures do only indicate the outcome of the perception of messages, we cannot rule out the possibility of negative tweets being more influential than positive tweets when informing posterior debate evaluations like the perceived performances of the candidates. Given the rather negative tone of social media discourses during media events, looking deeper into this puzzle seems to be a promising field for further research.

Although the presented study was carried out during an actual election campaign, built on a (back-then) recent and highly relevant case and only used authentic stimulus material, it obviously has some limitations. The subjects, even though they could participate in a non-laboratory setting, knew they participated in a scientific study. Furthermore, the participants only watched a short extract of debate rather than a 90 minutes debate at full length. These caveats clearly limit the generalizability of our findings. Despite these concerns about the external validity, we are convinced that more studies following a strict research design focused on internal validity are needed to understand the causal mechanism underlying possible second screening effects.

⁵ Also see Table 3.

References

- Anstead, N., & O'Loughlin, B. (2011). The Emerging Viewertariat and BBC Question Time. *The International Journal of Press/Politics*, 16, 440–462.
- Asch, S. E. (1951). Effects of group pressure upon the modification and distortion of judgment. In H. Guetzkow (Ed.), *Groups, leadership and men*. Pittsburgh, PA: Carnegie Press.
- Asch, S. E. (1955). Opinions and Social Pressure. *Scientific American*, 193, 31–35.
- Barnidge, M., Gil de Zúñiga, H., & Diehl, T. (2017). Second Screening and Political Persuasion on Social Media. *Journal of Broadcasting & Electronic Media*, 61, 309–331.
- Beckers, K., & Harder, R. A. (2016). “Twitter Just Exploded”. *Digital Journalism*, 4, 910–920.
- Chadwick, A. (2017). *The hybrid media system: Politics and power* (Second edition). *Oxford studies in digital politics*. New York: Oxford University Press.
- Chadwick, A., Dennis, J., & Smith, A. (2016). Politics in the age of Hybrid Media: Power, Systems, and Media Logics. In A. Bruns, G. Enli, E. Skogerbø, A. O. Larsson, & C. Christensen (Eds.), *Routledge companions. The Routledge companion to social media and politics* (pp. 7–22). New York, London: Routledge.
- Chadwick, A., O'Loughlin, B., & Vaccari, C. (2017). Why People Dual Screen Political Debates and Why It Matters for Democratic Engagement. *Journal of Broadcasting & Electronic Media*, 61, 220–239.
- Daschmann, G. (2000). Vox pop & polls: The impact of poll results and voter statements in the media on the perception of a climate of opinion. *International Journal of Public Opinion Research*, 12, 160–181.
- Davis, C. J., Bowers, J. S., & Memon, A. (2011). Social influence in televised election debates: a potential distortion of democracy. *PloS One*, 6, e18154.
- Deutsch, M., & Gerard, H. B. (1955). A study of normative and informational social influences upon individual judgment. *The Journal of Abnormal and Social Psychology*, 51, 629–636.
- Diakopoulos, N. A., & Shamma, D. A. (2010). Characterizing debate performance via aggregated twitter sentiment. In E. Mynatt, D. Schoner, G. Fitzpatrick, S. Hudson, K. Edwards, & T. Rodden (Eds.), *Conference proceedings and extended abstracts ; the 28th Annual CHI Conference on Human Factors in Computing Systems, April 10 - 15, 2010 in Atlanta, GA, USA* (p. 1195). New York, NY: ACM.
- Faas, T., Maier, J., & Maier, M. (2017). Einleitung - TV-Duelle als Forschungsgegenstand. In T. Faas, J. Maier, & M. Maier (Eds.), *Merkel gegen Steinbrück: Analysen zum TV-Duell vor der Bundestagswahl 2013* (pp. 1–9). Wiesbaden: Springer VS.
- Fein, S., Goethals, G. R., & Kugler, M. B. (2007). Social Influence on Political Judgments: The Case of Presidential Debates. *Political Psychology*, 28, 165–192.
- Freelon, D., & Karpf, D. (2014). Of big birds and bayonets: Hybrid Twitter interactivity in the 2012 Presidential debates. *Information, Communication & Society*, 18, 390–406.
- Gil de Zúñiga, H., Garcia-Perdomo, V., & McGregor, S. C. (2015). What Is Second Screening?: Exploring Motivations of Second Screen Use and Its Effect on Online Political Participation. *Journal of Communication*, 65, 793–815.

- Gil de Zúñiga, H., & Liu, J. H. (2017). Second Screening Politics in the Social Media Sphere: Advancing Research on Dual Screen Use in Political Communication with Evidence from 20 Countries. *Journal of Broadcasting & Electronic Media*, 61, 193–219.
- Gorkovenko, K., & Taylor, N. (2016). Politics at Home. In Unknown (Ed.), *Proceedings of the 9th Nordic Conference on Human-Computer Interaction - NordiCHI '16* (pp. 1–10). New York: ACM Press.
- Gottfried, J. A., Hardy, B. W., Holbert, R. L., Winneg, K. M., & Jamieson, K. H. (2017). The Changing Nature of Political Debate Consumption: Social Media, Multitasking, and Knowledge Acquisition. *Political Communication*, 34, 172–199.
- Grimme, C., Assenmacher, D., Preuss, M., Adam, L., & Lütke-Stokdiek, J. F. H. (2017). *Bundestagswahl 2017: Social-Media-Angriff auf das #kanzlerduell?* (PropStop Technischer Bericht).
- Gscheidle, C., Geese, S., & Gerhard, H. (2017). Berichterstattung zur Bundestagswahl 2017 aus Sicht der Zuschauer. *Media Perspektiven*, 2017, 594–606.
- Hahn, K. S., Lee, H.-Y., Ha, S., Jang, S., & Lee, J. (2017). The Influence of “Social Viewing” on Televised Debate Viewers’ Political Judgment. *Political Communication*, 42, 1–19.
- Houston, J. B., Hawthorne, J., Spialek, M. L., Greenwood, M., & McKinney, M. S. (2017). Tweeting During Presidential Debates: Effect on Candidate Evaluations and Debate Attitudes. *Argumentation and Advocacy*, 49, 301–311.
- Ito, T. A., Larsen, J. T., Smith, N., & Cacioppo, J. T. (1998). Negative information weighs more heavily on the brain: The negativity bias in evaluative categorizations. *Journal of Personality and Social Psychology*, 75, 887–900.
- Jungherr, A. (2015). Twitter use in election campaigns: A systematic literature review. *Journal of Information Technology & Politics*, 13, 72–91.
- Kätsyri, J., Kinnunen, T., Kusumoto, K., Oittinen, P., & Ravaja, N. (2016). Negativity Bias in Media Multitasking: The Effects of Negative Social Media Messages on Attention to Television News Broadcasts. *PloS One*, 11, e0153712.
- König, M., & König, W. (2017). Social TV: Die Twitter-Debatte zum TV-Duell: Kommunikation zum Hashtag #tvduell bei der Bundestagswahl 2017. *Media Perspektiven*, 2017, 630–638.
- Lang, G., & Lang, K. (1979). Immediate and Mediated Responses: First Debate. In S. Kraus (Ed.), *The great debates: Carter vs. Ford, 1976* (pp. 298–313). Bloomington: Indiana Univ. Press.
- Lee, E.-J., & Jang, Y. (2010). What Do Others’ Reactions to News on Internet Portal Sites Tell Us? Effects of Presentation Format and Readers’ Need for Cognition on Reality Perception. *Communication Research*, 37, 825–846.
- Leuchte, J. (2017). Interpersonale Kommunikation während und nach der Rezeption des Duells. In T. Faas, J. Maier, & M. Maier (Eds.), *SpringerLink : Bücher. Merkel gegen Steinbrück: Analysen zum TV-Duell vor der Bundestagswahl 2013* (pp. 173–189). Wiesbaden: Springer VS.
- Leviton, L. C., & Verhulst, B. (2016). Conformity in Groups: The Effects of Others’ Views on Expressed Attitudes and Attitude Change. *Political Behavior*, 38, 277–315.
- Lin, T., & Chiang, Y.-H. (2017). Dual Screening: Examining Social Predictors and Impact on Online and Offline Political Participation Among Taiwanese Internet Users. *Journal of Broadcasting & Electronic Media*, 61, 240–263.

- Lin, Y.-R., Keegan, B., Margolin, D., & Lazer, D. (2014). Rising tides or rising stars?: Dynamics of shared attention on Twitter during media events. *PloS One*, *9*, e94093.
- Maier, J., & Faas, T. (2003). The affected German voter: Televised debates, follow-up communication and candidate evaluations. *Communications*, *28*, 80.
- McKinney, M. S., Houston, J. B., & Hawthorne, J. (2014). Social Watching a 2012 Republican Presidential Primary Debate. *American Behavioral Scientist*, *58*, 556–573.
- McKinney, M. S., & Spialek, M. L. (2017). Political Debates. In M. Allen (Ed.), *The SAGE Encyclopedia of Communication Research Methods*. Thousand Oaks, California: Sage Publications.
- Mutz, D. C. (1998). *Impersonal influence: How perceptions of mass collectives affect political attitudes*. *Cambridge studies in public opinion and political psychology*. Cambridge: Cambridge University Press.
- Neuman, W. R. (2016). *The digital difference: Media technology and the theory of communication effects* W. Russell Neuman (1st Edition). Cambridge: Harvard University Press.
- Nyhuis, D., & Friedrich, J. (2017). Begleitung des TV-Duells auf Twitter. In T. Faas, J. Maier, & M. Maier (Eds.), *Merkel gegen Steinbrück: Analysen zum TV-Duell vor der Bundestagswahl 2013* (pp. 157–172). Wiesbaden: Springer VS.
- Rozin, P., & Royzman, E. B. (2001). Negativity Bias, Negativity Dominance, and Contagion. *Personality and Social Psychology Review*, *5*, 296–320.
- Schmehl, K. (2017). Diese geheimen Chats zeigen, wer hinter dem Meme-Angriff #Verräterduell aufs TV-Duell steckt. Retrieved from https://www.buzzfeed.com/de/karstenschmehl/willkommen-in-der-welt-von-discord-teil1?utm_term=.jeMkJp5Bo#.tnzWGDZE8
- Shah, D. V., Cappella, J. N., Neuman, W. R., Hanna, A., Bucy, E. P., Wells, C., & Quevedo, V. (2015). The Power of Television Images in a Social Media Age. *The ANNALS of the American Academy of Political and Social Science*, *659*, 225–245.
- Sherif, M. (1935). A study of some social factors in perception. *Archives of Psychology*, *27*.
- Thorson, E., Hawthorne, J., Swasy, A., & McKinney, M. S. (2015). Co-Viewing, Tweeting, and Facebooking the 2012 Presidential Debates. *Electronic News*, *9*, 195–214.
- Trilling, D. (2014). Two Different Debates?: Investigating the Relationship Between a Political Debate on TV and Simultaneous Comments on Twitter. *Social Science Computer Review*, *33*, 259–276.
- Tsfati, Y. (2003). Debating the Debate: The Impact of Exposure to Debate News Coverage and Its Interaction with Exposure to the Actual Debate. *Harvard International Journal of Press/Politics*, *8*, 70–86.
- Vaccari, C., Chadwick, A., & O’Loughlin, B. (2015). Dual Screening the Political: Media Events, Social Media, and Citizen Engagement. *Journal of Communication*, *65*, 1041–1061.
- Van Cauwenberge, A., Schaap, G., & van Roy, R. (2014). “TV no longer commands our full attention”: Effects of second-screen viewing and task relevance on cognitive load and learning from news. *Computers in Human Behavior*, *38*, 100–109.
- Von Hammerstein, K., Höfner, R., & Rosenbach, M. (2017). Wahlkampf mit Fake-News: Wie die Rechtsextremen mit Hetze mobil machen. Retrieved from

<http://www.spiegel.de/spiegel/wahlkampf-mit-trollen-wie-die-rechten-im-netz-mobil-machen-a-1167063.html>

- Waddell, T. F., & Bailey, E. (2017). Is Social Television the “Anti-Laugh Track?” Testing the Effect of Negative Comments and Canned Laughter on Comedy Reception. *Psychology of Popular Media Culture*. Advance online publication.
- Zerback, T., & Fawzi, N. (2017). Can online exemplars trigger a spiral of silence? Examining the effects of exemplar opinions on perceptions of public opinion and speaking out. *New Media & Society*, 19, 1034–1051.
- Zillmann, D. (2002). Exemplification Theory of Media Influence. In J. Bryant & D. Zillmann (Eds.), *LEA’s communication series. Media effects: Advances in theory and research* (2nd ed., pp. 19–42). Mahwah, NJ: Erlbaum.

Appendix

Table 5: Explaining performance evaluation and Perceived public opinion towards the candidates with alternative reference groups

	Results			
	Performance evaluation		Perceived opinion of others	
	Merkel	Schulz	Merkel	Schulz
Merkel Positive	0.428** (0.132)	0.009 (0.122)	0.414*** (0.107)	-0.076 (0.103)
Merkel Negative		0.077 (0.123)		0.146 (0.104)
Schulz Positive	0.185 (0.135)	0.250* (0.125)	0.179 (0.110)	0.356*** (0.106)
Schulz Negative	0.181 (0.128)		0.292** (0.104)	
Neutral	0.141 (0.130)	0.084 (0.121)	0.094 (0.106)	0.025 (0.102)
CDU/CSU	0.686*** (0.120)	-0.184 (0.116)	0.227* (0.098)	0.030 (0.098)
SPD	-0.208* (0.102)	0.850*** (0.098)	0.075 (0.084)	0.195* (0.084)
Constant	3.446*** (0.099)	3.051*** (0.088)	3.449*** (0.080)	3.035*** (0.074)
N	697	697	692	692
R-squared	0.070	0.120	0.033	0.035
Adj. R-squared	0.062	0.112	0.024	0.026

* p<0.05; ** p<0.01; *** p<0.001