The quality of political news in a changing media environment
Jacobi, C.A.

Citation for published version (APA):
Jacobi, C. A. (2016). The quality of political news in a changing media environment

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: http://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.
What do ongoing changes in the media environment, such as a popularization of news and a shift towards individualized online media, mean for political news quality? How can we define news quality in this context, and how can we measure it?

This dissertation argues that news quality standards should be more lenient towards characteristics of popular news, which may increase political interest and knowledge among specific audience groups. In the online environment characterized by information overload and vulnerable to bias as a consequence of audience click-rate tracking, drawing people's attention and keeping it is arguably just as important as transmitting political information.

Empirically, a content analysis of political news in Austrian online and print newspapers during the 2013 election campaign shows that in print, elite and popular newspapers perform rather similar across quality indicators, including content complexity. However, online, popular newspapers perform worse, even if we emphasize involvement alongside information.

Lastly, the dissertation discusses the benefits and limits of automatic content analysis methods, which are essential for measuring characteristics of political news content and quality in large datasets. It demonstrates the use of topic modelling, and compares its results to a previous study of frames in the news.
The Quality of Political News
in a Changing Media Environment
The Quality of Political News
in a Changing Media Environment

ACADEMISCH PROEFSCHRIFT
ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. D.C. van den Boom
ten overstaan van een door het College voor Promoties ingestelde
commissie,
in het openbaar te verdedigen in de Aula der Universiteit
op vrijdag 22 januari 2016, te 11:00 uur
door Carine Alida Jacobi
geboren te Leiderdorp
Promotiecommissie:
Promotor: Prof.dr. K. Schoenbach, Universiteit van Amsterdam
Copromotoren: Dr. W. van Atteveldt, Vrije Universiteit Amsterdam
Prof. dr. H.G. Boomgaard, Universität Wien

Overige leden: Prof. dr. mr. N. Helberger, Universiteit van Amsterdam
Prof. dr. J. Kleinnijenhuis, Vrije Universiteit Amsterdam
Prof. dr. P.C. Neijens, Universiteit van Amsterdam
Prof. dr. J. Peter, Universiteit van Amsterdam
Prof. dr. R. Vliegenthart, Universiteit van Amsterdam

Faculteit: Faculteit der Maatschappij- en Gedragswetenschappen
# Table of Contents

1. Introduction .................................................................................................................. 9  
   Aims and Research Questions ...................................................................................... 13  
   Case Study .................................................................................................................. 17  
   Contents of this Dissertation ...................................................................................... 20  
   References .................................................................................................................. 23

2. Evaluating the Quality of Political News in a Popularizing News Environment .................................................................................................................. 29  
   Introduction .................................................................................................................. 29  
   Journalistic Quality .................................................................................................... 30  
   (Political) News Quality According to Electoral Democratic Standards .................. 31  
   (Political) News Quality According to Participative Democratic Standards ........... 34  
   News Quality and Popular News .............................................................................. 36  
   Popular(-ized) News: The Negative View ................................................................. 39  
   Popular(-ized) News: The Positive View .................................................................. 44  
   Popularizing Electoral Democratic Quality Criteria .............................................. 48  
   ‘Popularized’ Electoral Democratic Quality Standards ........................................... 52  
   Popular(-ized) News and Electoral Democracy: A Research Agenda .................... 58  
   References .................................................................................................................. 60

3. Political News Content in Online and Print Newspapers ..................................... 71  
   Introduction .................................................................................................................. 72  
   News Quality from an Electoral Democratic Perspective ......................................... 74  
   The Quality of Political News in Online and Print Editions ..................................... 75  
   Research Questions and Hypotheses ........................................................................ 76  
   Case Study, Data, and Methods ................................................................................ 81
The Quality of Political News in a Changing Media Environment

1. Operationalization ................................................................. 82
2. Differences between Print and Online Editions ......................... 84
3. A Gap in Quality between Popular and Elite News Media? ....... 85
4. Discussion and Conclusion ..................................................... 89

References ....................................................................................... 93

4. Exploring News Content Complexity Using Network Analysis
............................................................................................................. 103

1. Introduction ............................................................................. 104
2. Complexity as a Quality Criterion for News Coverage ............ 105
3. Defining Content Complexity ................................................... 107

4. Indicators of Content Complexity Derived From Graph Theory...
   1. Interconnectedness .............................................................. 110
   2. Centralization ..................................................................... 112
   3. Heterogeneity ..................................................................... 113

5. Data and Methods .................................................................... 115
6. Operationalization .................................................................... 117
7. Results ..................................................................................... 121
   1. Interconnectedness .............................................................. 121
   2. Centralization ..................................................................... 124
   3. Heterogeneity and Relative Media Attention ....................... 125

8. Summary and Discussion ....................................................... 128
9. References ............................................................................... 132

5. Quantitative Analysis of Large Amounts of Journalistic Texts
   Using Topic Modelling ................................................................. 137

1. Introduction ............................................................................. 138
2. What is Topic Modelling? ....................................................... 139
3. How to Set Up an LDA Topic Model ........................................ 142
   1. Preprocessing: Tokenization, and Lemmatization ............... 143
   2. Feature Selection ............................................................... 143
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Choosing Parameters</td>
<td>144</td>
</tr>
<tr>
<td>Tool Support for LDA</td>
<td>148</td>
</tr>
<tr>
<td>Nuclear Topics: 1945 - 2013</td>
<td>148</td>
</tr>
<tr>
<td>Granularity: How Many Topics?</td>
<td>155</td>
</tr>
<tr>
<td>Discussion: Best Practices</td>
<td>157</td>
</tr>
<tr>
<td>References</td>
<td>161</td>
</tr>
<tr>
<td><strong>6. Discussion and Conclusion</strong></td>
<td>165</td>
</tr>
<tr>
<td>Summary</td>
<td>166</td>
</tr>
<tr>
<td>Discussion</td>
<td>170</td>
</tr>
<tr>
<td>Democratic Ideals</td>
<td>170</td>
</tr>
<tr>
<td>Quality in Elite and Popular News Media</td>
<td>171</td>
</tr>
<tr>
<td>Quality of Online News</td>
<td>173</td>
</tr>
<tr>
<td>New measures and Methods for Content Analysis</td>
<td>175</td>
</tr>
<tr>
<td>Avenues for Further Research</td>
<td>177</td>
</tr>
<tr>
<td>References</td>
<td>180</td>
</tr>
<tr>
<td><strong>Bibliography</strong></td>
<td>185</td>
</tr>
<tr>
<td><strong>Account of Funding and Co-Authors</strong></td>
<td>207</td>
</tr>
<tr>
<td>Funding</td>
<td>207</td>
</tr>
<tr>
<td>Contributions from Co-Authors</td>
<td>207</td>
</tr>
<tr>
<td><strong>Summary</strong></td>
<td>209</td>
</tr>
<tr>
<td><strong>Samenvatting (Dutch-language Summary)</strong></td>
<td>215</td>
</tr>
<tr>
<td><strong>Acknowledgements</strong></td>
<td>223</td>
</tr>
</tbody>
</table>
1. Introduction

For many people in modern Western democracies, mass media are an important source of information about what is going on in the world. Mass media help their audiences form an opinion on social issues, and give them advice on practical matters (Eide 1997; Connell 1998). This is why ‘quality’ journalism in mass media is considered important. But what is ‘quality’ journalism – does it only comprise the outcome of the journalistic process (news quality) or also its process? What exactly does it mean when news is called ‘high in quality’? How can we measure quality? And is there an essence of news quality that is valid across countries, media types (elite or popular, television or newspaper, online or offline), genres (political news, sports news, financial news) and time? Some quality criteria for news seem to apply universally, such as accuracy, transparence of the sources used, and independence (Shapiro 2010). Others are more important in some countries than others, such as separating facts and opinion (Schudson 1978; Schudson 2001), or in particular media types than others, such as a consideration of what the audience will like (Deuze 2005). Furthermore, quality criteria change over time (Picard 2014).

Within this broad universe of journalism types and qualities, this dissertation focuses on the quality of one particular type of journalistic performance: that of political news in print and online newspapers, during an election campaign. The term ‘political news’ as used here includes news about the activities of governments, political parties or other representatives, and their policies, as well as editorials, background reporting and the selection of letters to the editor printed in newspapers on this topic. In short, ‘political news’ contains all information about national politics that is offered by a particular news outlet during the campaign. ‘Quality’ for this type of news often has the connotation of ‘public value’, that is, news that makes a contribution to the greater good, to society as a whole (McQuail 1992), or more specifically, to a well-functioning democracy.
But how can political news contribute to democracy? This depends on what one considers the ideal form of democracy, and the role that news media should play within it. In the context of an election campaign, the electoral model of democracy (Sartori 1987) is often used as the ideal form of democracy. In such studies, quality criteria for political news are derived from the function that news media have within a democracy according to this model. According to the electoral democratic model, the main task of news media is to provide citizens with information about the policy plans and past performances of political parties and politicians (Strömbäck 2005). Citizens can then use this information to form an opinion on different political issues and candidates, and vote accordingly. Given the specific function of news media as information providers in a representative democracy within the electoral democratic model, news quality criteria derived from this model focus on characteristics of news stories that presumably strengthen this function. Such criteria include diversity, rational argumentation, relying on expertise, complexity, and a focus on issues in public policy rather than private events or personal characteristics of political actors (Sparks 2000; Jandura and Friedrich 2014).

However, news that adheres to these criteria is not necessarily very attractive to all audiences (Fiske 1990). This weakness of the notion of news quality derived from the electoral democratic model is especially relevant in the current online media environment characterized by individualization of the news offer and by an abundance of choice in (news) content. In such an environment it is more difficult to come into contact with political news at all if one is not particularly interested in politics. It is easy to just read something else instead, and if you do so often enough, your newsfeed will no longer show you news about politics at all. So even if all political news on offer would be high in quality, this does not help democracy if (groups of) people do not read or watch it. The current falling voting turnout in many Western democracies (Franklin 2004; Plasser and Ulram 2008) could be a sign that an increasing number of people do not find representative politics particularly interesting. This again raises the question if a concept of ‘political news quality’ that focuses
solely on the quality of the news on offer, but ignores those elements that make news attractive to read or watch, is suitable at all given the current media environment and problems with democracy.

A similar argument has been made before in studies on the role of popular news in democracy. While some see popular news as a threat to democracy because they fail to provide sufficient political information and focus on ordinary people and emotion rather than experts and ratio (e.g. Esser 1999; Sparks 2000) – thus following electoral democratic standards—, others recognize that popular news contributes to democracy in other ways. Rather than for the political information they provide, popular news media are valued for increasing political interest, and with that, political knowledge, among politically disinterested audience groups (e.g. Baum 2003; Baum and Jamieson 2006). As such, some scholars suggest using news quality criteria derived from the participative model of democracy rather than the electoral. The participative model requires increased political participation by everybody, no matter the level of expertise or mode of arguments – emotional or rational- used (Ferree et al. 2002). As such, news quality standards from this model could be used to evaluate the quality of popular news, according to some (Friedrich and Jandura 2012; Jandura and Friedrich 2014). A third model of democracy, the deliberative model, has recently gained popularity for evaluating the quality of political discussions online (e.g. Dahlberg 2001; Wright and Street 2007; see also Freedman 2000) again looking at more audience participation.

However, such studies have not led to a new set of news quality criteria that we can use to evaluate and compare the quality of political news in mass media of different types (popular, elite, etc.) and formats (print newspaper, website, etc.), in a specific context, such as an election campaign. Quality criteria derived from the electoral model seems to remain the standard for measuring political news quality in this context, despite their problems.

By this standard, the current media environment looks threatening for ‘quality’ news. The increasing competition among news outlets, their different funding structures and a shift towards online media
could lead to a deviation of all news from what is perceived to be high in quality (e.g. Blumler 2010). A lack of resources may force news outlets to cater to the tastes of an audience that is not really interested in serious politics (Sparks 2000, 3-4). And because the audience is not ‘trapped’ anymore by quality news such as public service newscasts (see e.g. Schoenbach and Lauf 2002) but highly fragmented, a growing share of citizens may encounter politics only through popularized news content, or avoid it altogether, resulting in a wider knowledge gap (Prior 2005). News sites that recommend particular news stories based on their popularity further increase this kind of selective exposure (Knobloch-Westerwick et al. 2005) and thus potentially contribute towards a spiral of news content lower and lower in quality, when held to electoral democratic standards. A popularizing and digitalizing media environment thus seems to pose new threats to the quality of political news, although one could question whether news media now serve democracy less than they did a few decades ago, or whether the quality criteria we use are simply no longer suitable.

But such questions of suitable criteria are not the only challenge journalism researchers face when dealing with the subject of news quality in the current media environment. Another challenge is more practical. I mentioned that one characteristic of this media environment is that there is more choice, that is, more media content (including news) on offer than a few decades ago. Moreover, people’s media diets consist of content from many different outlets (e.g. Trilling and Schoenbach 2013). As such, we need to include news stories from more different outlets, adding up to more news stories in total in order to analyze the quality of the news on offer in this fragmented media environment, or of the news actually watched or read by individual people. Also, the archives of online media provide great opportunities for studying changes in news (quality), but again, they tend to contain a lot of news stories. This means that we need methods that allow for greater quantities of data (texts) to be analyzed. Traditionally used methods such as manual content analysis might no longer be sufficient. Automatic content analysis methods are faster than traditional manual methods, which means that they can
analyze more texts. However, such methods require a different set of skills than currently available to most communication scholars, notably programming. Using insights on these methods from other disciplines for the study of news content and quality is another challenge for scholars who study this topic in our current context.

To sum up, we are faced with a changed and still changing news environment that may threaten the quality of news, if we hold on to traditional views of news quality. This has led to more discussion about the standards by which (political) news quality should be measured, to concerns about the quality of the (political) news currently on offer, and to new challenges regarding the methods to use in order to measure news quality. This dissertation aims to make a contribution to all three of these aspects: a discussion of quality criteria for political news that suit our current media environment, empirical findings on political news quality in different media outlets, and new methods and measures to measure it.

**Aims and Research Questions**

In the following chapters, I will first critically discuss the continued relevance of news quality criteria derived from the electoral democratic model, given the democratic benefits of popular news – which is low in quality following electoral democratic standards – demonstrated by several studies. I will give an overview of the dangers and benefits of popular news as its critics and supporters, respectively, see it. Discussing their arguments, I argue for a slightly changed set of quality criteria for political news during an election campaign; criteria which seem more suitable for the current media environment in which attracting people to actually read or watch political news in the first place might be a necessary aspect of quality. These criteria are a combination of those derived from the electoral model (with a focus on providing information) and from the participatory model (with a focus on making news interesting for its audience, and on increasing political participation). These ‘popularized’ electoral democratic news quality criteria, as I call them, form the theoretical contribution of the dissertation. They are meant to be relevant not only for academic research on news quality, for
news production by journalists, and for the general notion of news quality in society, which seems to be unduly focused on so-called quality newspapers. Recognizing the ‘quality’ aspects of popular news media as well, and, for newsmakers, making political news strong in such aspects, could help create better quality political news for everyone.

Empirically, I will focus on the question of how political news in popular media compares to that in elite media. I investigate how much political news is actually offered and how diverse this news offer is regarding the visibility of political actors, which are important aspects of political news quality. A major part of this study will address how ‘popularized’ news is, in terms of its emotionalization, focus on leaders and its simplification (which I measure as its opposite, complexity). Popular news characteristics are deemed non-desirable by electoral democratic standards, but are they really more strongly present in popular political news (see Connell 1998) during an election campaign? To find answers, I will compare political news in different media types – elite and popular newspapers, and online and offline news - for the aforementioned quality indicators (see Ferree et al. 2002; Jandura and Friedrich 2014). Are people who read their elite (broadsheet) newspaper online offered political news as high in quality as people who read the print version? And how different is the quality of political news in elite and popular papers?

To make a contribution to previous studies that answered similar questions, I focus on Austria, a country with an arguably more popularized and less polarized media market than the countries studied previously. Again, our view of popular news and its coverage of politics might need some nuance (both in the academic debate on political news in representative democracies and in society at large). Findings that support this form the empirical contribution made by this dissertation.

Lastly, I explore how new a number of new measures and methods can help to analyze news content and news quality given the current, changing media environment and the new quality criteria for political news that I argue we should employ. I introduce a measure for content complexity derived from network analysis, and explore how
to use text mining and topic modeling, two automatic content analysis methods, for the goal of measuring aspects of news content and quality. I will discuss the benefits and limits of these methods, so that other researchers may consider how such methods can be of use to them. This forms a third contribution, to content analysis methodology.

These aims lead to the following research questions:

**RQ1** Given recent findings and theoretical insights regarding the democratic functions of popular news, how relevant are quality criteria derived from the electoral model of democracy still?

**RQ2** What are the differences and communalities of political coverage in popular and in elite media in Austria, in terms of their (electoral democratic) quality?

**RQ3** How can new measures and (automatic content analysis) methods contribute to exploring news content and quality?

To answer these research questions, I will focus mostly on one case study, and use a combination of research methods.

A large part of this dissertation is based on a case study of the 2013 National Election Campaign in Austria. Austria is an interesting case due to its strong popular press, notably the so-called tabloid *Kronen Zeitung*, which has a market share of about 34% (Media Analyse 2013). Furthermore, Austrian elite newspapers are argued by some to be more popularized than popular newspapers (Magin and Stark 2014), suggesting an overall rather popularized press market. But at the same time, Austria prides itself of a strong public-service broadcasting system (Brüggemann et al. 2014), which indicates also a strong influence of public service quality standards in the overall news market. Practically, Austria has a strong online presence of newspapers and broadcasting, which have their content fully freely available. The final chapter does not focus on Austria. It is a methodological study, in which I explore the usefulness of the topic model Latent Dirichlet Allocation (LDA) (Blei, Ng and Jordan 2003) for the study of news content and quality. To this end, I partially
replicated an existing longitudinal study using American news coverage, and will discuss commonalities and differences.

This study is a multi-method one, combining new automatic content analysis methods that allow for quicker analysis of large datasets, with more fine-grained manual content analysis, as well as a thorough analysis of theoretical arguments and empirical findings from earlier studies. Because the research methods used are different for each chapter, a detailed description can be found there. A brief overview will suffice here.

- **RQ 1** will initially be answered using a categorization of earlier research (chapter 2). In the conclusion, I will come back to this answer given the empirical findings for RQ2.
- **RQ 2** will be answered using semantic network analysis and text mining, which are respectively a computer-assisted and an automatic, quantitative content analysis method. Semantic network analysis (Van Atteveldt 2008), in which trained student coders code news stories using an online content analysis toolkit (amcat.nl, see Van Atteveldt 2008), is used in chapter 4 for measuring news content complexity. This method displays the content of texts as networks of actors and issues, allowing me to use complexity measures derived from graph theory to explore the complexity of this content, and compare this between elite and popular papers. Text mining (Krippendorff 2013, 221) is used in chapter 3. This chapter compares political news quality in online and print editions on a number of indicators measured, using text mining (search strings). These indicators are all calculated from word occurrences, which can be measured with this method. Its speed (allowing me to include more news stories) and high reliability make this method preferable over manual or semi-automatic approaches.
- Lastly, **RQ 3** will be answered using the two methods discussed above (but focusing on their methodological aspects rather than the empirical findings they helped generate), as well as another automatic method, the topic model LDA mentioned earlier. Topic models cluster words based on patterns of co-occurrence in texts. They are potentially a great method for quickly gaining insight in
the contents (including quality aspects) of a large set of
documents such as news stories. The findings of the final chapter,
which focuses on the use of LDA for the study of news content and
quality, will provide the most important part of this answer. The
new measure for content complexity introduced in chapter 4, as
well as the benefits and limitations of using search strings for
measuring news quality indicators in chapter 3, will form the
other part.

Case Study

Before moving on to the substantive part of this thesis, it is relevant to
learn a bit more about the context of the findings presented there:
Austria, or specifically, its 2013 National Election Campaign. Although
I have already briefly touched upon some particularities of the
Austrian media system, a bit information could help the reader better
understand the meaning of the findings presented in this dissertation,
as well as the generalizability.

According to the categorization by Hallin and Mancini, Austria is a
democratic corporatist country (Hallin and Mancini 2004). Democratic
corporatist countries are characterized by high
newspaper circulation, a history of political parallelism in the press
but a diverse offer of newspapers, strong professionalization and self-
regulation of journalists, and strong state intervention, notably
through press subsidies and public service broadcasting (Hallin and
Mancini 2004, 67). A relatively weak media regulation, notably
regarding accountability and concentration, have led others to classify
Austria as a border-crosser between the democratic-corporatist and
the more politics-driven and less professionalized polarized pluralist
model (Karmasin et al. 2011). A recent empirical study by
Brüggemann et al. classified Austria differently altogether, showing
that its media system has more in common with Switzerland,
Germany and Great Britain – notably its strong public service
broadcasting – than with other Democratic Corporatist countries like
the Netherlands or Denmark (Brüggemann et al. 2014). Rather than
its exact classification however, the characteristics of Austria’s media
and political system that may help to explain the findings regarding news quality in this dissertation are more relevant here.

Firstly, for contextualizing news diversity, it is important to know about the specific political parties and ideologies important in Austrian society. In the early to mid-20th century, Austrian society was strongly divided between conservative Christians, voting for the ÖVP (Österreichische Volkspartei) and social democrats, voting for the SPÖ (Sozialdemokratische Partei Österreichs) (Plasser and Ulram 2003). The importance of these parties decreased since the 1960s, which with some delay also shows in their decreasing share of votes, e.g. in 1983, the ÖVP and SPÖ together gained 91% of all votes (Plasser and Ulram 2003, 27), against 58.5% of seats in Parliament at the time of the 2013 election campaign. As competing ideologies, the Austrian political landscape right before the 2013 election included two right-wing populist parties (FPÖ –Freiheitliche Partei Österreichs – and BZÖ – Bündnis Zukunft Österreich –; with respectively 21% and 6%), the Green party (11% of seats) and a number of smaller parties, notably successful newcomer Team Stronach, headed by a business magnate. News media thus have to balance between two very powerful established coalition parties, three opposition parties, and a number of newcomers.

Secondly, for contextualizing the importance of re-evaluating political news quality standards in times of news popularization, it is relevant to know a little about the Austrian media market, in particular the importance of both popular newspapers and public service broadcasting.

In general, Austrian journalists subscribe to what Hanitzsch et al. (2011) call’ traditional western journalism ideals’: the role of watchdog of the government, the goal of providing political information, but also of providing interesting news and motivating people to participate in political discussion and activity (Hanitzsch et al. 2011, 280-281). These ideals seem to fit both journalism that acts in the public interest and journalism that is made to be liked by its readers or viewers.
For newspapers, the Austrian market is characterized by high newspaper readership and strong press subsidies (Brüggemann et al. 2014), which makes it more likely that news media have resources available necessary for creating quality news. Press subsidies are given by the Austrian government in order to increase the plurality of available titles (see Murschetz 1998), despite which market concentration is extremely high (Seethaler and Melischek 2006). In absolute numbers, there are however many different types and individual titles of daily and weekly newspapers, including a considerable regional newspaper market (ibid.).

Popular and free newspapers in general and the tabloid newspaper *Kronen Zeitung* in particular have a very large market share in the Austrian newspaper market (Media Analyse 2013) and a high degree of perceived impact on politics (Plasser and Lengauer 2010, 93). Additionally, the media consumption of a considerable section of the Austrian population concentrates mainly on popular media, including the online version of popular newspapers (Trilling and Schoenbach 2013). Elite newspapers (*Der Standard, Die Presse* and *Salzburger Nachrichten*) are by comparison very small in circulation, which, as I noted previously, suggests a strong influence of popular news in the Austrian news market (see also Magin and Stark 2014). Since the mid-2000s, these paid newspapers are supplemented in some regions by free newspapers *Heute* and *Österreich*. Starting in 1995, Austrian newspapers also developed online editions, with variations between different titles regarding the degree of overlap with the print edition (García Avilés et al. 2009), but invariably offer all of the content online for free.

Furthermore, Austria has strong public service broadcasting, that is, it has a high market share and revenue (Brüggemann et al. 2014, 1051). Since news in public service media is regulated following quality norms (Costera Meijer 2005), this could provide a counterbalance to a strongly popularized news offer in newspapers – if the latter were to be actually low in quality. Public service broadcaster ORF takes care of all public service radio, television (on two channels), as well as a news web site. However, as the political parties are represented in its
Board of Trustees, there are some concerns regarding its independence from government (Lengauer 2006; Wenzel 2012). Austria was relatively late in its introduction of commercial TV channels; only in 2001 Austrian commercial channels were allowed on the market, although German commercial channels were widely watched in Austria for a longer time (Steinmaurer 2009). For news coverage on Austrian national politics, these channels will however be less relevant. Austrian commercial channels still do not have a strong market position (Karmasin et al. 2011). However, Austrian public service media are partly financed through advertising revenue, instead of through license fees only (Steinmaurer 2009), increasing their vulnerability to commercialization and its consequences for news quality (Breitenecker 2011).

Overall, the Austrian media system shows an interesting combination of on the one hand many (public) resources available for creating (political) news coverage, notably the support of newspapers and public service television by public means. The high readership of newspapers and high market share of public service television suggests that people may indeed actually read or watch the news supplied there. However, the Austrian media system seems also very popularized and somewhat vulnerable to the consequences of commercialization and political influence. This apparent contradiction provides an interesting case for studying the quality of political news content.

Contents of this Dissertation
As mentioned previously, in this dissertation, I look into the continuing relevance of news quality criteria derived from the electoral model of democracy. I compare how different media types (print and online versions of elite and popular newspapers) perform on these criteria using Austria as a case study. Lastly, I explore the use of digital methods for the study of news content and quality in large sets of news stories, as is required in a fragmented news environment. The dissertation consists of four chapters that contribute towards these goals. Below, I will briefly describe each chapter.
The second chapter, *Evaluating the Quality of Political News in a Popularizing News Environment*, is a categorization of theoretical arguments and empirical findings in earlier research that focuses on the concepts ‘news quality’ according to the electoral and the participatory model of democracy, and ‘popular(-ized) news’. In this chapter, I will explain the electoral and the participatory model of democracy, the role of media in democracy, and the criteria for news quality according to both respective models. Then, I will give an overview of previous studies that focus on the role of popular(-ized) news in democracy – news presented in such a way that it is attractive, interesting and understandable for a large audience. As I will show, popular news is evaluated negatively following electoral democratic standards, but positively by researchers whose notion of news quality is more in line with the standards derived from the participatory democratic model. I will then suggest a set of quality standards that combine criteria derived from both models, in order to give more credit to the role that popular(-ized) news can fulfill in democracy, and which I argue is especially important in the current media environment.

The third chapter and first empirical paper, *Political News Content in Online and Print Newspapers: Are Online Editions Better by Electoral Democratic Standards?* compares the news quality (according to the electoral democratic model) of Austrian online newspapers to that of their print version, using automatic content analysis (text mining using search strings). Websites have different format characteristics than printed paper, and the production process of online news is partially different from that of print news. Some of these differences may be beneficial for news quality (e.g. less limitations on space) while others may jeopardize it (e.g. things must be published quickly and cheaply). Also, websites provide new opportunities for adapting news to the likes and dislikes of the audience, by showing journalists and editors which stories are read most. As such, it is relevant to compare online and print news on a number of indicators for electoral democratic quality, to show how such differences may influence news content and style. The chapter makes this comparison both on the general level (online versus print) as well as on the level
of type (popular and elite newspapers). This approach allows us to see whether elite and popular newspapers exploit the opportunities of websites (for enhancing its quality and for popularizing it) in similar ways.

The fourth chapter, *Exploring News Content Complexity Using Network Analysis*, focuses on a specific aspect of news quality, complexity. High complexity is often seen as a quality criterion for news, although there is debate as to which forms of complexity are beneficial for which audiences. Still, measurement of complexity is often missing from studies on news quality or popular news. I suggest that this may be because of a lack of measures for content complexity (previous research mainly looked at textual complexity), which I contribute towards solving by introducing a new measure for content complexity. I use complexity indicators from network analysis (graph theory) to explore the complexity of manually coded political news content, which is displayed as networks of actors and issues. Also, the distribution of links within these networks is explored in order to explain why news in some outlets is more or less complex than that in others. I then compare elite and popular newspapers on these indicators, since elite papers are generally considered to be more complex than popular papers (e.g. Connell 1998), but it is not known whether this also applies to the complexity of their content, or only to their textual complexity.

In the current media environment, news quality is challenging to measure also because there is such a large quantity of news offered in this environment. Automatic text analysis methods that allow for faster coding and analysis of news stories can help researchers overcome this challenge. The final chapter, *Quantitative Analysis of Large Amounts of Journalistic Texts Using Topic Modeling*, explores the use of one of these methods, the topic modeling algorithm Latent Dirichlet Allocation (LDA) (Blei, Ng and Jordan 2003) for journalism (quality) research. LDA automatically calculates patterns of word co-occurrence in texts, and displays these as word clusters (topics) to be interpreted by the researcher. As such, it is a method for quickly finding out what large collections of texts ‘are about’ in terms of these topics. The question is however if topics can represent concepts that
are relevant to the study of news content and quality, such as (political) issues? This chapter answers this question by comparing the findings of an analysis conducted through LDA with an existing study. In order to do so, and at the same time use a dataset large enough to run an LDA model, I use a different case study than in the previous chapters: changes in American news coverage on nuclear technology since the Second World War. The goal of the chapter is however to investigate in which ways LDA can contribute to studying news content and quality, so the emphasis is on methods. The chapter describes the process of running an LDA analysis and interpreting its results, and makes recommendations on how to use this method for other researchers.

References


1. Introduction


1. Introduction


2. Evaluating the Quality of Political News in a Popularizing News Environment

Carina Jacobi, University of Vienna, Austria

Wouter van Atteveldt, VU University Amsterdam, The Netherlands

Klaus Schoenbach, Northwestern University in Qatar, Qatar

Abstract

Criteria derived from the electoral democratic model are still the norm in many studies on political news quality. Their suitability has been challenged by research that demonstrates the benefits of popular news for informing and involving politically disinterested audiences in democracy. The beneficial elements of popular and popularized news, however, would be called undesirable in the electoral democratic model. In this article, we evaluate the electoral democratic model and its counterpart, the participative one, as standards for high-quality news. We suggest a merger of the two models, by incorporating a number of ‘popularized’ political news quality criteria into the quality criteria derived from the electoral democratic model, and point out roads for further research in which these criteria can be used.

Introduction

Political news in mass media is often evaluated in terms of how much it contributes to democracy (Sparks 2000). However, there are different ideas among scholars regarding what ‘democracy’ means and what kind of contribution news media can and should make to it. The electoral democratic model (Sartori 1987; Strömbäck 2005) and its criteria for political news quality seem to be dominant in news quality research so far. It focuses on the purpose of news to inform
citizens about politics. Such news supposedly enables citizens to choose their political representatives rationally (Ferree et al. 2002), and thus sets particularly high standards for political news coverage during election campaigns.

The importance of this model is however being challenged by a different democratic model, participative democracy (Strömbäck 2005; Jandura and Friedrich 2014). This model requires citizens to be more actively involved in politics, and as such, has different criteria for news quality (Friedrich and Jandura 2012). This difference in criteria becomes especially clear in the discussion about the value of popular and popularized news (news made more attractive to a large audience through characteristics such as emotionalization, personalization and simplification).

In this article, we will first describe and discuss two concepts in more detail: news quality (journalistic quality in general, and political news quality by electoral and participative democratic standards, respectively), and popular(-ized) news. Then, we will typologize previous theoretical and empirical work on threats and benefits of popular(-ized) news for democracy. A critical evaluation follows, leading to an integrated concept of criteria for news quality in a democracy. We conclude the article with suggestions for news quality research that would follow from this concept.

**Journalistic Quality**

The concept of journalistic quality in general includes both norms for creating all types of journalism in a ‘good’ way, as well as norms for what the style and content of the outcome of this process – journalistic works – should look like. As others have noted, journalistic quality is hard to conceptualize, mostly because of a lack of clearly defined standards and agreed upon measures (Shapiro 2010). Journalistic quality is self-regulated by journalists and editors, and the skills of high quality articles or making excellent reportages is learnt on the job (Schultz 2007). These are said to be part of a journalistic ‘doxa’, an implicit set of (professional) values learned over time (*ibid.*). Also, there are variations in what is considered journalistic quality in different countries, where different ideas about
the societal function of journalism lead to different ideas about desired characteristics in excellent journalistic works. For example, should news stories only inform, or should they also advocate? (Esser 2008; Hanitzsch 2011). Formal evaluation of journalism is furthermore done by academics from various disciplines, who do not share the same concerns about its quality (Zelizer 2004).

Although it might be easy to come up with a vague idea of what journalistic quality entails, it is difficult to pin it down into a set of core criteria that are valid for all types of journalism in all contexts. Shapiro (2010) gives an overview of such core criteria, in an attempt to grasp what distinguishes ‘quality’ journalism from related forms of expression, such as blogs (Shapiro 2010, 145). Following the findings of a literature review of news quality as defined by different academic traditions, he suggested a number of standards, based on different key steps in the process of creating news stories. Quality journalism is independent, accurate, open to appraisal, edited, and uncensored. Independence refers to the journalist’s discovery of story topics and angles, which should not be influenced by personal connections. Accuracy refers to the gathering and examination of news facts, which should be done rigorously so as to make a description of reality which is as closely to the real world as possible. Open to appraisal refers to the interpretation of these gathered news facts: journalists should make clear which news facts lead to which choices in the way they tell the news story. Editing refers to the story’s style, which should be checked for clarity and readability by at least one editor. Lastly, the ‘uncensored’ norm means that quality journalism must be presented without influence from non-journalistic actors (Shapiro 2010, 155-157).

(Political) News Quality According to Electoral Democratic Standards

Because the criteria described above are only a minimal core set that apply to journalism in general, they do not ensure that news content contributes towards a well-functioning democracy, which is our concern here. If we want news to make a contribution to democracy (see Christians 2009), we thus need an additional set of criteria that
news has to adhere to in order to do so. Which specific quality criteria to use furthermore depends on the model of democracy one perceives as ideal. Indeed, democracy itself is perceived to be something different in different democratic models, from a basic procedure of government to follow from a collective spirit to be fostered in all citizens (for a detailed discussion, see Strömbäck 2005). The contribution that is asked from news media is obviously quite different, and smaller, in the former case than in the latter. Electoral democracy corresponds to standards of news quality commonly used in political news research (Jandura and Friedrich 2014).

Electoral democracy (Sartori 1987) is what Strömbäck calls a ‘realistic’ model of democracy (Strömback 2005, 334), resembling actual democracies (especially on the national level) in Western Europe. According to the ideal type of this democratic model, citizens choose representatives, and these representatives take the actual decisions regarding policy. For the contribution of news content to democracy, this means that citizens, as voters, need to be sufficiently informed about the policy plans of the different political actors, so they can vote based on which of these correspond most to their own views. Also, information about the past performances of political actors enables people to vote for another representative if they are not satisfied with these performances. It is one task of the media to supply people with this information, and to separate it clearly from other types of media content, such as entertainment, so that people can recognize it as relevant political information (Strömbäck 2005).

Models of the public sphere give additional criteria for news to contribute to a particular type of democracy. While models of democracy deal with different ways of how democracy can be organized and how news media should contribute to this, models of the public sphere (Habermas 2005) what kind of communication (who participates, and how should they communicate) can best contribute to this process (Ferree et al. 2002, 289). The model of electoral democracy is derived from the notion of a liberal-representative public sphere (e.g. Ferree et al. 2002; Martinsen 2009).
The liberal-representative notion of the public sphere extends the ideal of democracy by representation to the public sphere. In an extreme view, only expert representatives – be it politicians, scientists or interest groups – should participate in public debates, and communicate in a rational and detached way (Ferree et al. 2002, 292). Since the most important form of participation expected of ordinary citizens is voting, it is expected to them to follow an opinion on the proceedings in public debates, but not necessarily to participate themselves. The audience is invited to participate in public discourse only as long as it is well-informed, willing to listen and ready to be convinced by rational arguments. The rational, detached mode of communication in the public sphere is supposed to help people form an opinion, by giving them arguments for or against a particular policy plan, which may be persuasive or not (Jandura and Friedrich 2014, 353).

This process of democracy can be supported by media that comply with a number of quality criteria, which we refer to as ‘news quality derived from the electoral democratic model’. Although no consensus exists in the literature on what exactly these criteria are (see e.g. Ferree et al. 2012, 316; Martinsen 2009, 60), it is possible to distinguish some core aspects that are essential to ‘quality’ political coverage in an electoral democracy. Jandura and Friedrich (2014) summarize these broadly as relevance, diversity, and professionalism (Jandura and Friedrich 2014, 354).

Relevance limits coverage to those actors and issues that matter to the democratic process: the policy plans of politicians and political parties. Furthermore, their past performances are also important; performances both in terms of policy and in terms of how well politicians function professionally, so that citizens can vote them out of office if they are not satisfied with these performances in the past (Strömbäck 2005).

Diversity refers to the degree to which different ‘ideas’ are represented (e.g. McDonald and Dimmick 2003) in political news content. Diversity is a broad norm that can be used to evaluate the performance of media systems or news outlets (Napoli 1999). It
originates in John Stuart Mill’s concept of a ‘marketplace of ideas’ – the idea that it is beneficial for democracy when people learn about a multiplicity of voices or opinions on a variety of issues, because it is likely that the best idea will win in the end (Mill 2011 [1859]; Voakes et al. 1996, 582). In the electoral democratic model, diversity calls for coverage that covers the full range of these actors and issues, that is, political representatives and other voices that play a role in decision-making, as well as the full range of issues that are involved.

Lastly, professionalism includes a number of content and style characteristics essential to good political coverage (which happen to be quite similar to some of the general news quality criteria we discussed earlier): factuality (information that is accurate, up to date, and complete) and impartiality (balanced and neutral) for content, and detachment and civility for style (Jandura and Friedrich 2014, 354-355). Detachment – coverage that is not emotionally involved, but distanced from politics – is promoted by a strict separation of information and entertainment, and an unemotional style in informative news stories (Friedrich and Jandura 2012, 410). In addition to these criteria, there is also the implicit and very basic criterion that there should be a sufficient amount of news about politics available at all in news media (even popular outlets), without which it would be impossible media to fulfill their task of informing citizens about politics (Sparks 2000, 28).

(Political) News Quality According to Participative Democratic Standards

Electoral democracy is not the only model used as a basis for the evaluation of (political) news content – it does not fit all actual democratic contexts (direct democratic campaigns, for example). Furthermore, some critics of the electoral democratic model and its quality standards argue that this model is too elitist, since it prefers it if those who are not willing or able to become informed and argue rationally do not participate in public discussions, including the media. Instead, they suggest participative democracy with its participative public sphere is the preferred model to use (Friedrich and Jandura 2012; Jandura and Friedrich 2014).
In a participative public sphere, the ideal is to maximize the participation of citizens in political decision-making (Ferree et al. 2002, 295). Rather than just being well informed and vote, it requires people to take actively part in political discussions. Furthermore, it emphasizes that the essence of democracy is an attitude within citizens that incorporates involvement with society by ordinary people on every level, from the grass-roots decision-making in local communities to national and international politics (Strömback 2005; Martinsen 2009). Participation should not be limited to voting only: people themselves, rather than their representatives, should be able to influence policy directly. News media should reinforce this attitude and provide people with “the kind of knowledge and information that facilitates collective action, participation and engagement” (Strömbäck 2005, 336). The reason for this emphasis on high involvement is the distrust of power inherent in this model, as well as a view of society as constant conflict between groups different in power status. These differences in status will be reproduced if people do not participate in politics themselves (Ferree et al. 2002, 297).

Because of this emphasis on participation instead of representation, and on participation by everyone, not just the well-informed, the participative model has different standards for quality in political news content. By its standards, simplification, emotionalized discourse and even incivility are encouraged because of their mobilizing potential. In a radical view of participatory democracy, “slogan-based politics” – politics communicated through strongly simplified and emotionalized statements – would be preferred over informed debate, if this increases participation (Feree et al. 2002, 298). A more moderate view emphasizes feelings of trust and bonding towards politicians and civic associations (Strömbäck 2005, 336) as motivators of societal participation. The participatory model furthermore favors news content that connect social issues to people’s lifeworld, that portrays people as active citizens rather than passive victims, and asks from newsmakers to include people more into the process of creating news as well (ibid., 340).
News Quality and Popular News

As the above description suggest, news in different types of news media will perform better when held to the standards of different models. The electoral democratic model, which still seems dominant in political communication studies that focus on news content, favors the type of (political) news provided by ‘elite’ or ‘quality’ media (Friedrich and Jandura 2012), notably newspapers: media that target a relatively highly educated and politically interested audience. That political news in these media is presumably high in quality by (electoral-)democratic standards is a somewhat circular argument, since providing ‘quality’ news, including a focus on political issues (e.g. Sparks 2000), is part of what defines these media, at least in the everyday usage of the term. However, here, our focus is on popular and popular(-izing) news, which, by this logic, is lower in quality. Some argue that popular (political) news would perform better when held to the quality standards derived from the participative democratic model (Friedrich and Jandura 2012). But before discussing the quality of popular news, let us step back and consider what defines and characterizes popular news in the first place.

As a starting point for an overall definition of popular news, Dahlgren’s distinction between the thematic and the pragmatic dimension of journalism is useful (Dahlgren 2000). The thematic dimension of journalism refers to its content – to how journalism represents reality. The pragmatic dimension refers to journalism as communication, to the techniques it uses to attract an audience and make the news story interesting and understandable for them (Dahlgren 2000, 50). So this dimension does not necessarily refer to following audience preferences, but to striving to create news stories that will be paid attention to and that are well understood by the intended audience. ‘The popular’, then, “is that which enhances the pragmatic dimension for the majority” (Dahlgren 2000, 51), in other words, a focus on those characteristics of news stories that make them attractive, interesting and understandable for a large audience. Rather than simply considering ‘the majority’, which sounds like a large homogeneous audience, it is perhaps better to think of the audience for popular news as a large group of people that even could
be diverse in terms of educational background, profession, place of living, and gender (for Austria, see Trilling and Schoenbach 2013). This focus on making news attractive, interesting and understandable, in principle for everybody, is arguably the defining characteristic of popular (-ized) news as opposed to elite news.

But how can we recognize this characteristic of popular news empirically, in the content and style of its political coverage? Emphasizing the pragmatic dimension can be done in various ways. This results in a number of empirical characteristics that are often considered typical for popular news. These characteristics can be evaluated positively or negatively, depending on one’s view on what political news should look like, and on what one thinks the contribution of (political) news to society and to democracy should be. Here, we will describe some characteristics that are often considered to be typical for popular news, and that are relevant in regard to its potential democratic contribution (see also the overview by Reinemann et al. 2012).

A central characteristic of popular news compared to elite news in general is that it has less attention for public life and for ‘hard news’ issues (national and foreign politics, and economics) and more for private life and for ‘soft news’ issues (lifestyle, celebrities, and entertainment) (Sparks 2000). For political news in particular, this characteristic means that popular news media will probably have less of it than elite news media. Furthermore, the focus on the private life typical for popular news manifests itself as a particular genre of news: human interest stories, news stories about the experiences of unknown individuals (e.g. Hayashi 2000; Fine and White 2002) that elite media supposedly have less of. Another genre typical for popular news is ‘service journalism’ – stories that provide readers all kinds of advice (Eide 1997), which again can be seen as part of this private focus of popular news. Lastly, simple, everyday language rather than complex discourse is typical for popular news (Connell 1998). Overall, the emphasis is on what the audience likes to read or see, and on what is perceived to be directly useful for the audience (Brooks 2000) rather than useful for a greater good.
Within *political* news, too, popular news media have a focus on the private over the public. Typical are news stories about the private life of politicians, also called privatized coverage, (Van Aelst et al. 2012). This is however not the case in all countries: some news cultures, such as Germany, keep the private lives of politicians out of the news (Klein 2000). Furthermore, popular news, both overall and news on politics specifically, uses a style that emphasizes emotions (emotionalization) or ‘the senses’ (as opposed to ratio) overall (Grabe et al. 2001). Some see especially negative emotions as typical for popular political news, which is sometimes referred to as scandalization (Esser 1999).

There is some disagreement among researchers about which of these characteristics are central to popular news, that define it empirically, so to say. Some emphasize its focus on ordinary people, and define popular news as ‘the people’s news’ – news about issues that are put on the news agenda by the people, directly or through journalists (e.g. Hayashi 2000; Örnebring and Jönsson 2004). To others, privatization and scandalization are the main characteristics of popular news (e.g. Esser 1999, Donsbach and Büttner 2005). For the discussion about the critical view and the positive view of popular news that follows, it is not necessary to pin down exactly what popular news is and what it is not, but to understand which of its characteristics are criticized and/or praised for their potential effects on democracy. Before moving on to this discussion, it is first important to go into one more common perception of popular news, namely that there is more and more of it, due to a commercializing and digitalizing media market.

‘Making news attractive to a large audience’ and ‘making news that sells well’ are often considered to be the same (Costera Meijer 2005). As such, increasing competition among media outlets, which is said to force these outlets to focus on a larger audience, is seen as a cause of increasing popularization or ‘tabloidization’ (e.g. Rowe 2011). That is, more news outlets use popular news elements in the news they create, and they use them to an increasing extent. A number of longitudinal changes in the news market have presumably led to more competition between news outlets. Newspaper circulation is falling, especially since the introduction and progress of online news
media (Sparks 2003; Westlund and Färdigh 2011). The growth of online news media is also relevant for news popularization, since the Internet gives media outlets more possibilities to trace people’s preferences through their clicks on particular individual news stories, leading to more pressure on journalists to create media content that is popular with large audiences (Anderson 2011).

There are two scenarios for the effects of the overall increased competition in the news market, to which the audience-tracking contributes, on news popularization. In one scenario, the possibility to trace audience preferences on an individual news story level furthers popularization even in elite media, because all audiences prefer popular(-ized) news. In the Netherlands, journalists admit that they change the content and placement of news stories based on what people click on most (Welbers et al. in press). For newspapers, this information about audience preferences obtained through clicks in the online version is also used for changing the placement of news stories in the print version. So print newspapers, too, are not immune to the popularization impulse given by the Internet (ibid.). In the other scenario, different audiences are seen to have different tastes, which means that commercialization may lead to a greater divide in the market between elite and popular news (e.g. Sparks 2003; Curran 2005), when both types cater to the tastes of their respective audiences.

**Popular(-ized) News: The Negative View**

But why would following audience preferences be a threat to democracy, and as such, be a sign of bad quality, by standards derived from the electoral democratic model? Critics of popular news media consider this a problem because it means prioritizing news about topics that people find interesting over news that is relevant using standards based on the public interest (e.g. Rooney 2000; Curran 2005), which broadly overlap with the relevance criteria for political news posed by the electoral democratic model. The fear of these critics is that prioritizing what the audience likes leads to less political news in popular media than they consider ideal. Some popular news media carry so little information on political issues and performances
of political actors that people who rely on these media only do not get the chance to be sufficiently informed about politics. As such, they cannot vote in the thoughtful and informed way preferred by the electoral democratic model; they might not know whom to vote for at all (e.g. Rooney 2000; Sparks 2000). This would mean that popular news hampers democracy.

But also in the political coverage of popular media, their critics see some problems. Following electoral democratic standards, political news should focus on the plans of political parties regarding various issues, and on the performances of politicians. This is the information that should matter for people’s voting decisions, according to the electoral democratic ideal (Strömbäck 2005). News about the private life of politicians, one type of popular political news, is not seen as relevant, and could encourage people to take this information into account when they vote. Personalized stories are also seen as harmful because of their potential negative effects on democracy, notably news stories about the private life of politicians (Van Aelst et al. 2012) which are said to increase cynicism (Jebril et al. 2013).

Lastly, the style of popular news – its presumed negativity, emotionalization and scandalization (e.g. Donsbach and Büttner 2005) are seen as a problem, for four reasons. Firstly, these characteristics are the direct opposite of the detached and civilized style that is seen as ideal in the liberal-representative public sphere (Jandura and Friedrich 2014, 356). Again, this could jeopardize the ideal way of voting according to this model, because people are presented with news that encourages them to form a political opinion based on their emotional reaction, rather than on what they should consider rationally to be the best party or candidate.

Secondly, journalists working for popular news outlets select those events that can be turned into emotionalized human-interest stories or scandals, using these characteristics as news values (Skovsgaard 2014) rather than the political relevance criteria posed by the electoral democratic model. This means that the popular news audience might miss out on information on important political issues.
Thirdly, this could turn into a cycle of selective exposure: people prefer the negative and scandalized aspects of politics presented to them by popular media, and therefore popular news media report politics in this way, thus confirming people’s views that politics only has this side (see Stroud 2008). This would decrease news diversity—politicians not involved in scandals are not considered newsworthy by (popular) news media, which thus does not offer their audience the chance to learn about politicians who are not involved in scandals. This could lead to a misinformed (and not to mention cynical) audience (see also Cappella and Jamieson 1997).

Fourthly, populist parties averse to traditional elite politics are said to systematically benefit from this cycle of scandalization (Plasser and Ulram 2003; see however also Walgrave and de Swert 2004; Boomgaarden and Vliegenthart 2007). This would mean that also the impartiality of journalism is harmed, since popular news media employ a style that confirms the worldview of some parties more than that of others.

Aside from these negative consequences for democracy of popular news as a category, critics of popular news also fear popularization—the increase of popular news elements in all media—especially in online news. Earlier, we mentioned two scenarios: all news becomes more popularized, or popular and elite news grow further apart. Critics of popular news see negative consequences of both scenarios.

In the popularization-scenario, critics fear that over time, the audience loses its ‘tolerance’ for detached, analytical news and their attention spans become too short to process long and complex ‘quality’ journalism. This is complemented by the fact that in an online environment, traditional news media also compete with newer media such as news blogs written by citizens, who do not necessarily follow general journalistic quality standards. This may lead to the lowering of such standards in traditional media as well (Tumber and Waisbord 2004, 1147). Lower journalistic quality in general might also mean lower news quality by electoral democratic standards, since the two types of quality overlap partially, as we mentioned previously. If all news then becomes worse in quality by electoral democratic
standards, the quality of democracy itself is at stake. Then, people are no longer provided with the information to vote based on issues, but will likely vote based on personality traits of candidates, or other information that the electoral democratic model considers irrelevant, if they vote at all. Consequently, those in power can pursue their own preferred policy without democratic control (Curran 2005, 140).

The second scenario sees the online media market as a force that magnifies different audience interests and, as a result, differences in knowledge: the politically interested will have access to more political news and thus become better informed as they read or watch this news. Those without much political interest, however, will become even less knowledgeable, since the popular media they prefer will offer ever less political news, knowing that political stories get fewer audience clicks (Sparks 2003) and thus bring in less advertising money. As such, the chance of becoming at least somewhat politically informed while not being interested in politics decreases. The personalized selection of news on social media newsfeeds especially makes it less likely that these media bring an audience with low political interest into contact with political news (the so-called ‘trap effect’) compared to television or newspapers (Schoenbach and Lauf 2004, 179).

Popular outlets are accused of focusing on celebrity and entertainment ever more strongly, and do so at the expense of their only characteristics sometimes prized by its critics, namely that they put citizens’ (political) issues on the news agenda (Hayashi 2000). Instead, they use voices of ordinary people to confirm the editorial opinion advocated by the outlet itself rather than letting people speak for themselves (Örnebring 2006), and address people as individuals rather than as a collective (Conboy and Steel 2010). This again decreases the potential for popular news to function as a kind of ‘news of the people’ with its own political discussions. News in popular outlets might thus be further ‘popularizing’ (or ‘tabloidizing’) too, due to increased competition and new technologies (Hayashi 2000; Uribe and Gunter 2004). By doing so, it moves further away from elite news in terms of quality. And as a consequence, the knowledge gap (Tichenor et al. 1970) might increase.
A final potential consequence of the greater choice and personalized news recommendation in online news media is increased selective exposure to one’s preferred political ideology (Garrett 2009; Messing and Westwood 2012). While not directly related to news popularization, this is also a consequence of news media that follow audience preferences more closely, and make the media environment more fragmented as a result; this time by political ideology rather than by level of political informedness. This threatens the diversity within news outlets (Goldman and Mutz 2011).

Empirical studies provide some evidence for some of the negative consequences of popular(-ized) news for democracy. A preference for ‘soft’ popular news is associated with low political knowledge (Prior 2003). Exposure to broadsheet newspapers and public service media leads to more political knowledge and also a higher motivation to vote, while exposure to outlets with less political content does not (De Vreese and Boomgaarden 2006). Studies suggest that the way people vote becomes less issue-based when the news does not cover issues (Takens et al. 2015), while issue-based voting is considered ideal according to the electoral democratic model. There is also some indication that privatization increases cynicism among voters (Jebril et al. 2013), which may leave people unmotivated to vote at all.

As for the changes in the degree of popularization overall, both the scenario of general popularization and the scenario of increased fragmentation due to increased popularization in popular outlets have some empirical evidence (Welbers et al. in press; and Sparks 2003, respectively; see also Reinemann et al. 2012), but from different national contexts and thus different media markets. There seems to be no general trend of either increasing popularization or increased fragmentation online.

To summarize, critics accuse popular news of not informing people sufficiently about politics, and of informing people in the wrong way, by focusing on those aspects of politics that are not relevant to an issue-based voting choice, and by using a style that encourages people to form an emotion-based opinion rather than in a rational-detached
fashion. As such, it does not contribute to democracy according to the ideals of the electoral democratic model, and is thus low in quality by the standards of this model. At its worst, popular news has characteristics that even discourage people to fulfill their civic duty (voting) altogether.

This is a rather harsh view of popular news. It furthermore suggests that a large section of the population is currently under- or misinformed (the large audience for popular news media), while only a small elite is not (the audience for elite media). Only in countries where public service broadcasting succeeds in informing a broader audience about politics, this would presumably not be the case. This view furthermore assumes that ‘the people’ are not of themselves interested in issue-based news (only the elite is), but that they should be, and that journalists should not decrease their interest in politics by using emotional, engaged reporting or depicting politicians as private people. Furthermore, critics fear that media in general are taking the popularization route, or that one group of media does so while another does not, increasing fragmentation of the audience.

**Popular(-ized) News: The Positive View**

The positive view of popular(-ized) news more or less agrees with the negative view about the news focus and style that are typical for popular news, but argues this focus and style have positive effects on democracy. This positive view is partly a defence against the attack on popular news by political communication scholars, of scholars who take a more audience-focused research approach. Some of these scholars suggest using quality standards derived from the participative democratic model instead of the electoral, for all news media.

Firstly, popular news is praised for its potential to involve ‘the people’ in politics *because* it focuses on ordinary citizens and their experiences and on political leaders (Mazzoleni 2000), portraying them ‘as human beings’ (Klein 2000), and *because* it uses everyday language and an emotionalized or sensational style to present these stories (Grabe et al. 2001). The positive camp postulates that these characteristics are not harmful, as critics say, but that they help
people to process information, and attract them to this information in the first place (Eilders 2006). Already in 1937, Hughes stated that human interest stories inform people about what is happening in society, not “the political news” (Hughes 1937, 73). More recently, scholars assume that mixing politics and popular culture brings politics closer to the people, especially to disinterested groups (e.g. Van Zoonen 1998; Bird 2000). Rational, unemotional coverage might not be the most effective for mobilizing people to participate in politics (Meijer 2001), which popular journalists also have in mind when writing personalized stories (Skovsgaard 2014).

A further advantage of popular news pointed out by its supporters is that it includes a different kind of information about representational politics than elite newspapers, which again increases audiences’ understanding of political issues. Popular newspapers offer people advice on how to pursue their interests within the political realm, which is part of the aforementioned ‘service journalism’ (Eide 1997). Also, policy consequences on the daily life of ordinary people are often illustrated by human interest stories that are useful to people because it speaks more to their own life than the abstract description of policy that elite media offer (Brookes 2000). In other words, popular news outlets use the language and life world of common people to deal with larger issues (Gripsrud 2000; Hayashi 2000), which is said to work better for conveying information than just describing the larger issues per se.

A slightly different view does not value the characteristics of popular news explicitly, but does note the potential of popular media as bundles of discourses, some serious and political, and some entertaining and non-political (Sparks 2000). Entertaining stories may attract people, but because they receive news as a bundle, they will also at least see some political news, and perhaps this will interest them enough to pay attention to this particular news item, or read the news story that incidentally caught their attention (Sparks 2000)—the so-called trap effect mentioned earlier (Schoenbach and Lauf 2002; Schoenbach and Lauf 2004).
Perhaps because the positive view of popular news does not regard it as a threat to democracy, researchers who adhere to this view do not look into questions of popularization of elite media and its consequences, but instead emphasize the continuity between types of popular news. Popular news has always been there (see e.g. Schudson 1978) and will continue to fulfill different societal functions than elite media (Curran 2005, 139). Rather than arguing for different functions of popular media, some scholars see its benefits mainly in its presentation, and argue that ideally, popular and elite audience groups would receive political information that is similar in its topics, but adapted to the group regarding its style or level of complexity (Dahlgren 2000; Gans 2009).

A second, stronger view takes the audience-focused approach a step further. Instead of saying that popular news has less news about politics (and so does not well on the ‘relevance’ criterion of the electoral democratic model), supporters of this view argue that the understanding of ‘politics’ or ‘public issues’ as ‘representational politics’ and ‘issues in public policy’, respectively, are too narrow. Media content, including entertainment and other non-news, can also influence people’s opinion on political issues and morality (Bird 2000; Moy et al. 2014) and therefore have a political dimension. Some issues are not yet part of the political debate, but are problems in the daily lives of groups of citizens, and therefore they are political (Hayashi 2000). News coverage should also include these issues, rather than just the issues put onto the agenda by established politicians. This can make these politicians, as well as other officials or other people (more) aware of such issues, so that they can undertake action (ibid.). Örnebring and Jönsson call tabloid newspapers an alternative public sphere for this reason: they can function as a bridge between people who put new issues onto the agenda, and these can then be picked up by more elite media and reach greater society and political actors in this way (Örnebring and Jönsson 2004). In a similar vein, Gans (1979) argues that news should represent more perspectives than just the views of and issues held important by those in power.
Empirical studies on popular television journalism have shown some support for the idea that popular news can make disinterested groups more interested in politics (Baum and Jamieson 2006) and increase people’s understanding of politics in ways that elite news cannot (Baum 2003). The ‘trap effect’ has been shown to work for television—it transmits political information to a large audience, including those who would not explicitly seek out political information in other media (Schoenbach and Lauf 2002). These studies thus suggest that popular news, especially on television, includes non-elite audiences into the public sphere, while news with the style and focus common in elite media would reinforce their exclusion.

To summarize, the positive view of popular news prioritizes ‘what works for people’ rather than ‘what would be ideal in theory’ – it favors those news styles that make people pay attention to news content, and that increase the understanding of politics. In this way, popular news can contribute to democracy. If all news would be like elite news high in quality according to electoral democratic standards, it would exclude a large group of people from political discussions. Arguably, this is an important concern, given current problems of growing disinterest in democracies. Popular news could actually help—which is exactly the opposite of what the negative view of popular news regards its influence to be. Even though popular news is low in quality by standards derived from the electoral democratic model, it can still make a democratic contribution.

However, then why not call such elements aspects of quality? Supporters of the stronger view discussed above seem to do so. They argue that news quality criteria derived from the electoral democratic model are not democratic enough, because they are not open enough to the needs of their audiences. They challenge what is considered political, arguing that the definition of ‘the political’ should come from the people. And they praise popular news media for providing a way for people to indirectly influence the political agenda themselves. Online popular media have great potential for doing so, due to their interactive features enabling citizens and representatives to
communicate directly with each other. These characteristics, together with those described in the previous paragraph, makes popular news potentially high in quality by standards derived from the participative model of democracy. This would mean that ‘quality’ popular news and ‘quality’ elite news contribute to different forms of democracy.

**Popularizing Electoral Democratic Quality Criteria**

As the above made clear, the negative view of popular news adheres strictly to electoral democratic criteria for news quality, and mainly criticizes those aspects of popular news that do not adhere to these criteria. Supporters of popular news instead seem to be supportive of any style element or news topic (such as the private lives of politicians) that the audience favors and that therefore may increase motivation to participate or that may increase understanding. Coming from a line of research that focuses on news coverage during election campaigns and typically uses news quality criteria derived from the electoral democratic model, we argue that these criteria neglect the important qualities of popular news that its supporters emphasize. Not only within the participatory democratic model, but for electoral democracy too, it is important to also look at the political news that different audience groups actually read or watch, and the information that they retain from it. The electoral democratic model focuses too much on news as supply, and ignores the needs of different audience groups in terms of motivation and understanding. Even if all supplied news was high in quality by the standards of the electoral democratic model, this would not assure that all audience groups would benefit from it to the same extent.

Should we then switch to quality criteria from the participative model, favoured by supporters of popular news, in order to appreciate the contributions of both elite and popular news? We argue that doing so would miss the point, since different types of democracy as they exist in reality ask for different theoretical models in evaluating how they function, which parts (such as the contribution made by news media) of the overall system can be improved, and how this can be done. It is thus important to look at how particular news
2. Evaluating the Quality of Political News

Quality criteria contribute to a well-functioning democracy according to the model used, whether this is electoral, participative or yet another form of democracy.

For news quality criteria derived from the electoral democratic model, we argue that their classical form does not achieve this, given the beneficial effects on democracy that studies on popular news show. We should ‘popularize’ these criteria, so that political news that is interesting and understandable for its audience (which ‘quality’ news is apparently not, for certain audiences), and that is thus more likely to attract these audiences to political news and as such, inform them about politics, is also considered a sign of quality. Currently, many Western democracies struggle with falling election turnouts (Franklin 2004), especially among young people (Goerres 2007), while the online media environment makes it more difficult for the politically uninterested to encounter political information (Schoenbach and Lauf 2004). As such, making it more likely that these audience groups, too, will read or watch at least some political news is important, and those elements of political news that make this more likely should be considered elements of quality – for the media that target these audiences.

Also, one could agree with the supporters of popular news who favor the participative model that the electoral democratic model discourages the participation of the uninformed and of non-rational, meaning that the outcome of debates will tend to reflect the interests of the current elite (Curran 2005, 139). Still, as Curran argues, without political information and debate in the media, citizens will become passive about issues, and no longer hold the government accountable (Curran 2005, 140). Diverse political information is thus still essential, but the focus of this information and its style should be adapted to the needs of different audience groups, so that their motivation to be informed and vote accordingly is maximized. Doing so, we can formulate quality standards that are not automatically dismissive of popular news, and thus fairer than the ‘classical’ electoral democratic standard.
Firstly, while the provision of factual and detached political information by news media is important, its necessity for supporting democracy is perhaps overemphasized in studies that take electoral democracy as its norm. Given the volatility and the lower voting turnout in a number of current Western democracies, the contribution that news media could make beyond just providing information should be giving more attention. This contribution includes involving people more into politics, guiding them towards relevant information (Macdonald 2000), and helping them build an opinion despite an overload of news and other types of information. In this era of an abundance of available information and media outlets, enlarging the chance of selective exposure (Knobloch-Westerwick et al. 2005), it cannot be taken for granted that audiences read the news that would be important for them by electoral democratic standards, that is, news that contains information about societal problems and the views on political actors on these problems. Information should thus be packaged in such a way that makes it likely to be read and understood by its intended audience; providing facts alone is not good enough. Therefore, journalism that is a little bit more ‘geared to the impact of the text’ rather than (just) its content, as Costera Meijer (2001, 190) describes popular news, can be beneficial for elite news as well, as long as is taken into consideration that different styles may have different impacts on different audiences.

Furthermore, the involvement and understanding required by the participative model is relevant for the electoral model as well. Although an involved audience is typically not considered to be the priority for this type of democracy (Strömbäck 2005), electoral democracy cannot function when people do not vote. Popular news shows that increasing involvement can be done in a way that solves another problem with the quality standards derived from the electoral democratic model – that it leads to political news that has a bias against ordinary people. ‘Elite’ media overlook “social lives, social histories, ordinary individuals, and the activities of day-to-day life” as Picard (2014, 504) nicely sums up. Popular media, if their description by their supporters is correct, do not have this problem. Furthermore,
human interest stories are what people like to read (Bird 2000). If such stories can help explain political issues in ways that people can identify with, that motivate them to vote or find out more about the issue and help them form an opinion, then why would they necessarily be ‘bad’ quality? The same argument applies to the use of emotions. Perhaps it would thus be better to focus on distinguishing the different ways in which emotions can be used in (political) news, and find out the conditions (particular types of emotionalized discourses, media for particular audience groups, particular topics) in which emotionalized coverage can help democracy.

Also, not all audiences have the time and ability to process a large quantity of political information every day. As such, rather than striving for as much political news as possible, it is more important that the news agenda is reasonably diverse (so that people still learn about different views even without having to devote much time to the news) and, to avoid fragmentation, not too dissimilar between outlets (Gans 2009). Additional depth and complexity should be available to those who are interested, but should not be seen as a requirement, and ‘more political news’ should not automatically be seen as ‘better’. The challenge for popular journalists then becomes how to capture the essence of a complex political problem or discussion and explain it in a clear and attractive way (Gans 2009).

Popularizing electoral democracy theoretically might also help against a problem inherent in this model. This problem is electoral democracy’s limited capacity for encouraging political change: it benefits established parties and elite views. This can be addressed by changing the norms for relevance and diversity somewhat, so that it is considered good if news media give new parties and alternative views more attention, and thus allow these to become more prominent in society.

Overall, we argue that even though electoral democratic news standards may be problematic because they set high requirements that are incompatible with certain aspects of popular news, their core is still important in certain democratic contexts, such as national election campaigns in representative democracies. In a democracy-
by-representation, people need information on how both everyday
life problems are greater societal issues are linked to the policy plans
of politicians, and also on the success or failure of politicians to
execute their policy plans, or other aspects of their public office – both
for established parties and small or marginal ones. It is however not
only this information that counts in quality news, but it also requires
a style that help people connect to it. As such, the detached, complex
and abstract style and focus on the status quo preferred by the
electoral democratic model might not be as relevant or as beneficial
as once thought, and that should be reconsidered, especially given the
changing news environment and current democratic problems.

‘Popularized’ Electoral Democratic Quality Standards

At the beginning of the article, we discussed three aspects of quality
political news derived from the electoral democratic model: news
should be relevant, diverse, and professional—which in this context
means factual, impartial, detached and civil (Jandura and Friedrich
2014). Even if we ‘popularize’ quality criteria for political news, that
is, we incorporate the strengths of popular news within the concept of
‘quality’, these aspects are still relevant. However, what they
specifically and subsequently mean, as well as how they should be
measured, changes slightly. Below, we will give an overview of those
empirical indicators of news quality that we think have to be
interpreted and measured differently, and discuss how we should
‘popularize’ these indicators.

1. Amount of Political News

Firstly, as mentioned above, if democracy requires informed citizens
and it is a task of the news media to inform people, there has to be a
substantial amount of news related to political actors and issues in
order to fulfill this purpose. Of course, this criterion immediately
evokes a number of questions, such as ‘When is news political?’, ‘How
much political news is enough?’, and ‘Are the answers to these
questions the same for every audience?’. To answer the first question,
we need a way to distinguish the political from the non-political. ‘The
process of deciding about all affairs public’ is a working definition of
politics. All news related to actors taking part in this process or affairs (issues) being decided upon is ‘political news’.

As to the second question, we cannot point to any research that answers this. The amount of political news is often measured relatively, as share of the full content of a newspaper or newscast. The higher this proportion, the more strongly the outlet is associated with being an ‘elite’ or ‘quality’ medium, and the lower the proportion, the stronger the association with ‘popular’ or ‘tabloid’ (e.g. Uribe and Gunter 2004). So, it seems that the more an outlet focuses on politics, the better in quality it is regarded to be – although hypothetically, a 30-page tabloid could in absolute terms still contain more political coverage than a 14-page up-market free daily.

The electoral democratic model asks media to supply citizens with sufficient political information. It seems to assume, though, that there should not be any difference in the supply for different audiences. In reality, not everybody has the same amount of time per day, the same amount of interest and capability to be willing and able to process political news. There may be a minimal amount of news, e.g. during an election campaign, to at least make people aware that they are supposed to vote. More important, however, is that whatever news there is should be interesting and relevant enough for people to actually read and watch it.

2. Relevance

So, the second quality indicator concerns the relevance criterion: when does (political) news contribute to a well-functioning democracy, and when does it not? According to the electoral democratic model, citizens are expected to have opinions about social issues, knowledge about who is in power and about their record of past performances, as well as about the policy plans of different parties (Strömbäck 2005, 337). News that contributes to this is seen as relevant. This basic notion of relevance can still be maintained for our ‘popularized’ quality criteria, but we will expand it somewhat.
Information on the private life of politicians, which is typical for popular news, as we discussed earlier, has a problematic status: it could of course be somewhat related to politicians’ performances as representatives of the people, but may not directly influence their professional functioning as good politicians. Then, it should not matter. So, arguably, information about the private life of office holders, or polls and discussions of media performance are less relevant for meeting the expectations that the electoral democratic model has of citizens. However, because such information still has to do with political actors, and might contribute to their overall evaluation (for example to their perceived trustworthiness and the consistency between the values they preach and those they actually practice), it would be hard to dismiss it as completely non-political. Privatized news may also make people more interested in politics – perhaps it draws their attention to news stories that they otherwise would not read, and that also include information on issues or on the professional performance of politicians.

The relevance criterion for news could thus be broadened—at least for certain audience groups— with particular types of political coverage that were not considered relevant by traditional electoral democratic standards, if these can demonstrably contribute to political interest or understanding.

3. Diversity

The electoral democratic model and its liberal-representative public sphere focus on politics-by-representation. The voices or opinions of different politicians or other representatives are thus especially relevant for diverse news. Issue diversity and the range of opinions about the importance of these issues and how they should be solved are other types of diversity relevant for representing different ‘ideas’ in the news. However, too much diversity might perhaps decrease the focus necessary for public discussion. Fragmentation could be a consequence: it becomes easy for particular audiences to select only news stories with those views that suit their own. (Issue) diversity is thus more of a balancing act rather than the more, the better: journalists should learn the importance of issues for their audiences
in order to focus on the most relevant ones, but not to such an extent that the agenda of some news media becomes one-sided or completely different from that of others.

Furthermore, for political parties, the electoral democratic model is said to favor a reflective norm when measuring diversity (Ferree et al, 2002, 292): the coverage of political actors should be proportional to their power status (Van Cuilenburg 1997). In practice, power status of parties is often operationalized as their number of seats in parliament (McQuail 1992, 201). The danger, of course, is that existing power relations are reinforced, and that new voices do not get a chance. During election campaigns especially, consequences of following the reflective norm are that new parties get no media attention at all, and that small parties too will barely have the chance to inform voters of their plans through the media (Jandura and Friedrich 2014, 356). A norm that favors an equal amount of media attention for every category (Van Cuilenburg 1997) seems more suitable for creating a ‘marketplace of ideas’: it is more impartial and helps avoid (partisan) selective exposure (Goldman and Mutz 2011). Again however, such a norm may result in ‘too much’ diversity at the expense of focus, and in addition seems not very feasible when there are many different political parties participating in a campaign. A balance between openness to new voices on the one hand and relevance (to be determined by newsmakers taking into account the opinions of their audiences) thus seems the most suitable.

Another element of diversity concerns politicians within parties. Representative politics is carried out by a large number of political actors. Policy plans can be presented as being the plans of the entire party (Rahat and Sheafer 2007, 66), but within that party too, individual representatives may have their own foci or even oppose the party line, and may even campaign for themselves. In many Western democracies, citizens can (also) vote for individual candidates. Political coverage that accurately reflects political reality would thus have to show different levels of the political process: the general plans and record of the party, of the top candidate, as well as of other candidates. Admittedly, this ideal is (too) complex.
However, campaign coverage often focuses on top candidates, a phenomenon known as presidentialization (Langer 2007; Van Aelst et al. 2012). Audiences seem to prefer to read about (famous, powerful) people. So for media outlets that need to sell their news, this focus makes sense (Mazzoleni 2000). But it actually means a diversion from the accurate depiction of political reality preferred by the electoral democratic norm because it tends to over-represent the importance of political leaders in politics (Takens et al. 2015).

Despite these dangers of a strong focus on leaders, especially when combined with an extensive discussion of their private life, it seems to have benefits when the news uses leaders or top candidates as representatives for political parties. Because it obviously helps people identify policies with faces and could thus support understanding, memorization, and political motivation.

4. Complexity

Complexity can refer both to news style (longer sentences, more difficult words) and content (more interrelations between different concepts). A simple style is considered to be typical for popular newspapers (Connell 1998) and seems to make news more attractive and easier to grasp. In most studies of news quality or popular news however, complexity is not taken into consideration. Given the function that it may have for understandability, we think that a suitable level of style and content complexity should be included as aspects of news quality (see also Gans 2009).

Content complexity is a new concept that we propose here (but see Suedfeld and Tetlock 1977; Kleinnijenhuis et al. 2015 for related concepts) refers to how complexly politics is represented in the news: how many different issues and actors are involved, and how often are different actors and issues connected to each other in the same news story. This latter aspect distinguishes it from diversity. We assume that politics is complex, both as a process (e.g. political debates) as well as in terms of the social problems and issues it addresses in policies, and their potential consequences. One could thus argue that quality news reflects this complexity, and thus gives citizens better insight in the realities of political debates or of social issues. As with
2. Evaluating the Quality of Political News

the amount of political news however, some people might be more willing and able to digest complex coverage than others

Rather high complexity per se, it is thus more important is thus that the level of complexity offered by the news matches the level which the audience can appreciate, both regarding content and style (length of sentences, grammatical complexity, etc.).

5. Detachment

According to the electoral democratic model, political coverage has to provide citizens with political information that helps to form their own opinion through a rational process. This means that political coverage should be restrained in its use of emotions, and instead present and compare different opinions in a rational, detached way (Martinsen 2009).

However, especially in current society where emotions take more and more center stage (Richards and Rees 2011), detached coverage may be unattractive to news audiences. In a news environment where choices on what to read or watch are abundant, news audiences can select scandalized, emotionalized celebrity news stories any time over detached political coverage. Even though it might be better in theory to keep emotions out of politics and out of political coverage as much as possible because they detract from making well-informed judgements (Kuklinski et al. 1991), the consequences of adhering to this ideal (groups of people who turn away from political altogether) might be even worse for democracy than letting it go to some extent. So instead of dismissing emotionalized coverage, it is perhaps better to evaluate how emotions could be an aspect of news stories that both attract and inform or explain. The use of emotions in this way should then be seen as a sign of quality.

Overall, there is a basic dilemma between attractive and functional news that has to be balanced: News may be very attractive to its readers or viewers but not contribute to democracy, and the other way around. So, emotions fulfill a dangerous role: one the hand, they make news more interesting, on the other hand, they may mislead their audience. This also applies to diversity, complexity, and the
amount of news: in democratic theory, it seems that the more, the better, but this overload could have negative consequences for the necessary focus and news understanding of audiences with low political interest. Therefore, it is important for all these criteria to continue studying how they work in news as communication, meaning how they contribute or detract from attracting and informing specific audience groups. Especially for lesser-interested audiences, we think it is better not to use the perfectly informed citizen as a standard, but to focus on the difference compared to the worst-case scenario: people who turn away from political news altogether, because they cannot understand or identify with it.

**Popular(-ized) News and Electoral Democracy: A Research Agenda**

What, then, should empirical research on political news focus on? As discussed previously, earlier work on news quality by electoral democratic standards often sees a binary distinction between popular and elite (‘quality’) media, where the former stands for ‘bad quality’ and the latter for ‘good quality’. We argued that we should be more lenient regarding the amount of information, and an emotionalized and simplified style, if such characteristics help to attain better understanding and more involvement. This argument has consequences for research on news quality. First and foremost, research into political news quality should also include reception studies, rather than only considering news supply, to sharpen insights into which styles and foci work best for motivating and informing different groups of people, and on different platforms (television, online, print). This includes also further investigating the possible link between popular news and the increasing support populist parties mentioned briefly in the discussion of the negative view of popular news.

However, this does not mean that content analysis in itself is not important. Especially relevant is the possibility that existing studies may have contributed to a rather black and white view of popular versus ‘quality’ news, whereas the difference between the two may not be so big (see e.g. Magin and Stark 2014), or either group is
2. Evaluating the Quality of Political News

perhaps not so homogeneous (for example, Takens et al. 2010 show higher diversity in free dailies, which could be considered a popular type of newspaper, than in paid popular newspapers). Below, we will list a number of ways in which previous content analytic research on popular news and political news quality can be extended in order to address this problem.

Most importantly, news quality according to electoral democratic standards should be measured in a broad selection of popular and elite outlets, and include online outlets of various types (online editions of popular and elite newspapers, citizen blogs, online-only news sites, etc.). The aim of this is to gain insight into what the actual differences between these outlets are on the different indicators of news quality. Previous research on popular(-ized) news and news quality tends to compare only outlets of one type over time, or only compare one elite versus one popular outlet, which is not sufficient for getting a more nuanced view of news quality in popular news. Online outlets, too, are often considered together as one group, but here too, it would be better to distinguish between popular-oriented and elite-oriented outlets (see Dimitrova et al. 2014). Aside from giving a more nuanced view of how various outlets score on quality indicators and how this relates to popularization, this type of broad comparative content analysis is a prerequisite step before doing audience studies – first, one has to see what is there, before investigating whether it works as communication.

Secondly, much of the existing knowledge on popular(-ized) news is based on British, American and German tabloid newspapers and infotainment television, which serve as a kind of ideal type of popular news. These cannot simply be generalized to other countries, as international differences such as cultural values, but even something seemingly unrelated such as laws that protect the privacy of politicians, may lead to international differences in popular news characteristics (Esser 1999; see also Gripsrud 1992; Sparks and Tulloch 2000). Even if popular news in those countries does not do well according to electoral democratic quality standards, perhaps it does in other countries that have a media system that is less polarized
between elite and quality, such as the Netherlands (Deuze 2005) or Austria (Magin and Stark 2014).

Thirdly, much of the existing research looks at coverage in general instead of at the (from a democratic perspective) especially relevant case of political coverage, and often considers routine times rather than election campaigns. Election campaigns are especially important for electoral democracies, because it is then that news can influence people’s vote decisions most directly, especially for late deciders. Whether news in general, including lifestyle news or the entertainment section, are depoliticized, emotionalized or in another way low in quality does not matter so much for democracy, but it would be more problematic – by electoral democratic standards—if this were the case for political news during an election campaign. This specific type of news during this specific time thus needs to be investigated separately.

Fourthly, certain characteristics of popular news, specifically the low complexity of their content, have not been included in news quality research very much, even though these characteristics might be relevant aspects of quality, especially from an audience perspective. We should find such characteristics, not only from theory but also inductively, and integrate them into our understanding of news quality.

As such, to sum up, research on popular(-ized) political news during election campaign using different countries and a selection of outlets broader than the typical tabloid and elite newspapers or infotainment television programs are required to give a more complete and nuanced picture of what popular news and the popularization of news, specifically for political news in an electoral democracy, are about. Findings regarding their (differences in) content can then be compared to traditional and ‘popularized’ electoral democratic criteria, and in a next step, be used in reception studies.

References

Anderson, CW (2011) Between Creative and Quantified Audiences: Web Metrics and Changing Patterns of Newswork in Local US
Evaluating the Quality of Political News


The Quality of Political News in a Changing Media Environment


Grabe, Maria E., Zhou, Shuhua, and Barnett, Brooke (2001). Explicating Sensationalism in Television News: Content and the Bells and...
The Quality of Political News in a Changing Media Environment


Welbers, Kasper, Van Atteveldt, Wouter, Kleinnijenhuis, Jan, Ruigrok, Nel, and Schaper, Joep (In press) News selection in the digital era: professional norms vs. real-time audience metrics. Accepted for publication in *Journalism*.


3. Political News Content in Online and Print Newspapers

Carina Jacobi, University of Vienna

Katharina Kleinen – von Königslöw, University of Zürich

Nel Ruigrok, Erasmus University Rotterdam

Abstract

The electoral model of democracy holds the ideal of citizens who are well informed about political issues and actors, and regards it as a task of news media to provide citizens with high quality information. Against this ideal, the quality of political news in online news outlets is highly contested. While pessimists point out the dangers of increased competition for quality news online, optimists emphasize the potential benefits of unlimited space and interactivity online. To see which view holds true, this article compares political news in popular and elite print newspapers and their respective online editions, during the 2013 National Election Campaign in Austria. Findings show that online editions score better than paper editions regarding the amount of political news, (party) diversity, and emotionalization, but differences between newspaper types were notable. Whereas elite newspapers cover politics online more extensively than in print, the reverse is true for popular newspapers. Leader focus is also strong in popular papers online. We conclude that the gap in quality between political news in

1 This chapter is derived from an article published in Digital Journalism, 21 September 2015, copyright Taylor & Francis, available online: http://www.tandfonline.com/doi/10.1080/21670811.2015.1087810.
elite and in popular newspapers is larger online. This might contribute to a wider gap between a well-informed elite audience and a lesser-informed popular news audience, when audiences switch from print to online news.

Introduction

It has been two decades since print newspapers entered the online news market with their own online version. In the meantime, many other online news media have sprung up and compete with online newspapers both for readers and advertising revenue (Humprecht and Büchel 2013). This increase in competition, combined with the faster cycle of online news, has led to fears of negative consequences for political news quantity and quality in online media (e.g. Plasser 2005). Online news sites are increasingly used as a news source, especially by young people (e.g. Mitchelstein and Boczkowski 2010). This means that a poorer political news offer in these media could contribute to the problems of low political interest and knowledge that exist in many Western democracies (Sparks 2003; Lee 2007; De Waal and Schoenbach 2010).

A normative perspective favouring electoral democracy (Sartori 1987) stresses the importance of political coverage that sufficiently informs people about politics, and presents a realistic picture of political events (Strömbäck 2005). As such, political news high in quality may help people to vote ‘correctly’, that is, for the political party that closest represents their own opinion (see Lau and Redlawsk 1997; Kovach and Rosenstiel 2007; Wahl-Jorgensen and Hanitzsch 2009), while low-quality news that provides an incorrect or incomplete picture of politics may hamper their ability to do so. Aside from other influences such as personal interest and historical allegiance, news media are an important influence on how people vote (Kiousis et al. 2006; Kleinnijenhuis et al. 2007). Media provide up-to-date information about politicians' performances and viewpoints, and often reach large audiences. Especially during election campaigns, it is thus important that (political) news is high in
quality, in order to be most beneficial to the electorate, and as such, to democracy.

In this paper, we will compare political news during an election campaign in online and print newspapers on a number of indicators related to news quality from an electoral democratic perspective. Although the electoral democratic model is a strongly idealized and, for news, supply-focused model, it provides clear criteria for the information that news should supply during campaigns, which have often been used in research on media quality and commercialization or popularization (see e.g. Reinemann et al. 2012). Other functions of news media in democracy, or criteria of news quality related to its reception – that is, how well certain characteristics of news content aid in improving attention to the news and understanding of the news by its audience (Dahlgren 2000) – can and should be studied separately, using different criteria (Jandura and Friedrich 2014). Here, we look at news supply using indicators related to the information function of political news in a democracy, and the degree to which this coverage is characterized by a commercial logic (see Landerer 2013). Commercialism often conflicts with criteria for quality derived from the electoral democratic model. Moreover, we assess differences between so-called elite and popular newspapers, since popular newspapers are presumed to be influenced more strongly by a commercial logic than elite newspapers, and are subsequently seen as harmful for democracy by some (Sparks 2000; Rooney 2000; Skovsgaard 2014). Our overall research question is as follows:

**RQ** How do the online editions of popular and elite newspapers compare to the print editions regarding the quality of political news during an election campaign, following electoral democratic quality standards?

We answer this question using political news coverage from six Austrian national newspapers (three popular, three elite; in both their online and print editions) during the Austrian National Election campaign of 2013. We focus on newspapers and their online editions
because of their traditional function as provider of political news during campaigns, in addition to television. Austria is an interesting case because of its overall rather popularized press market (Magin and Stark 2014). Most research on quality in online and print newspapers has been conducted in the United Kingdom and Germany, where elite and popular media may be further apart in terms of quality (Magin and Stark 2014; see also Sparks 2000). By comparing print newspapers with their online equivalents for both popular and elite newspapers, we shed light on the aspects of quality of political news coverage (during an election campaign) held important by the electoral democratic model in today’s changing media landscape.

**News Quality from an Electoral Democratic Perspective**

What is ‘quality’ news? Aside from a general set of criteria and good journalistic practices, such as accuracy and transparency regarding sources (Shapiro 2010), the quality of particular genres of journalism depends on the purpose attributed to these genres. For political news, the general expectation is that news contributes to democracy (Christians 2009). In turn, how it should do so depends on what form of democracy one regards as ideal.

The electoral model of democracy (Sartori 1987) is one of these forms. It is what Strömbäck (2005, 334) calls a “realistic” model of democracy, closer to actual democracies in Western Europe than other democratic models such as the participative or the deliberative model, which require much more involvement by citizens. In the electoral model, citizens choose their representatives, but otherwise their role in politics is limited. Citizens do not, and should not, influence policy directly (Sartori 1987). This type of democracy thus means that people have to know whom to vote for in order to have their viewpoints represented in parliament. Also, according to the electoral model, they should ideally vote based on these viewpoints, or for politicians that they think perform well as representatives (e.g. Ferree et al. 2002; Jandura and Friedrich 2014).

News media have the task of providing voters with sufficient and correct information about politics so that they are able to vote in an informed way (Strömbäck 2005; Lau and Redlawsk 1997), while
avoiding any type of discourse that may distract from forming issue-based opinions. This task forms the basis of a number of quality criteria for political news during an election campaign. Firstly, the news should provide enough information about politics. This means not only that there should be a large quantity of news about politics, but also that it should include the viewpoints and performances of all political parties taking part in the election, without bias toward a particular party. A strong focus on leaders at the expense of other candidates (Takens et al. 2015) or coverage with strong emotional rather than detached overtones (Martinsen 2009) could hamper this ideal of rational opinion-building and voting.

While elite media supposedly adhere to electoral democratic quality standards (Friedrich and Jandura 2012, 404), popular or tabloid news is often regarded as the opposite of quality news, by being apolitical, trivial, simplified, emotionalized and sensationalized (e.g. Esser 1999; Sparks 2000; Rooney 2000; Bakker and Scholten 2013). Their focus on what the target audience wants to read, and thus on what sells (to this audience directly or to advertisers) is seen as the cause for these undesirable characteristics; a focus confirmed by popular journalists (Skovsgaard 2014). The same line of argument also applies to online newspapers, which can follow their audiences’ preferences through tracking their clicks on individual news stories. This could make even ‘quality’ media more likely to publish what their audience wants to read rather than what they should know according to democratic ideals (Welbers et al., in press).

The Quality of Political News in Online and Print Editions
However, the discussion of online media and their contribution to or risk for electoral democracy is more complicated than just this general popularization or commercialization argument. Compared to print newspapers, online editions have a number of characteristics relevant in this context.

Firstly, the size of the online edition is less limited by its format (a website) than print news (a paper with about the same amount of pages every day). Also, the role of the audience is more pronounced, both directly in the comment sections of stories and discussion fora
(e.g. Thurman 2008), as well as indirectly through click rates. This latter aspect influences news selection, not only for the online edition (Vu 2014), but also for the printed newspaper (Welbers et al., in press).

Finally, online news can be published anytime, and can always be updated or removed later (Deuze and Yeshua 2001). In practice, this flexibility puts journalists under pressure to publish news quickly, in order to be the first to publish certain news stories (Anderson 2011), while fact checking can be postponed. This time pressure furthermore seems to promote practices such as copying press releases and wire-service content or using the content of other media as a source (Boczkowski and De Santos 2007; Boczkowski 2009).

Given these characteristics, the online editions of elite newspapers might widen or narrow the gap between them and their popular counterparts. This gap might widen, for instance, because popular newspapers may use the Internet to produce news to pursue their commercial goals even more strongly while elite newspapers may use it to produce high-quality news, as happened a decade ago in the United Kingdom (Sparks 2003). On the other hand, the gap could also narrow because quality outlets may give in to the presumable forces of the market, especially online, where less money can be made from subscriptions (unless paywalls are used, which is not the case in Austria). Elite newspapers in both Sweden and The Netherlands, for example, ‘tabloidize’ in their online edition and become more like their popular counterparts (Andersson 2013; Welbers et al. in press; see also Reinemann et al. 2012).

Below we discuss the consequences these differences may have for news quality, following our criteria for political news quality according to the electoral model of democracy.

**Research Questions and Hypotheses**

*Amount of political news*

A first indicator for the democratic quality of news during an election campaign is the amount of news about politics, since it is assumed that more supply of political news and thus political information
3. Political News Content in Online and Print Newspapers

provided by the media will lead to better informed citizens (e.g. Sparks 2000; McLachlan and Golding 2000). With the exception of election campaigns, popular newspapers presumably devote a low percentage of their total coverage to hard news topics such as national politics, while much of their coverage is on non-political soft news topics (Rooney 2000). Because of the large space available on websites compared to in print newspapers, we would expect that the amount of political news online will generally be higher than in print. However, resources and time are still limited even though online media provide more space (Oschatz et al. 2014). Moreover, one could also argue that the interactive elements of websites make it easier for journalists to focus on the news that audiences want to read (Vu 2014)—which is presumably soft news rather than political coverage (Welbers et al., in press). Our first sub-question thus reads:

**RQ1a What is the amount of political coverage in online and print newspapers?**

Given the presumed focus on hard news in elite newspapers (Sparks 2000), we expect outlets from this type to contain a larger amount of political news than popular newspapers. However, it is possible that this difference is larger, or instead smaller, in online editions, if one media type has different priorities or strategies than the other. This leads to the following question:

**RQ1b Is there a gap between the amount of political news in popular and in elite newspapers, and is this gap larger or smaller online compared to print?**

*Diversity*

A second requirement of news media in an electoral democracy is diversity (e.g., McQuail 1992; Mutz and Young 2011). This aspect of news quality is often associated with the ideal of the public sphere as a ‘marketplace of ideas’ (e.g. Voakes et al. 1996, 582; Napoli 1999, 8; Mutz and Young 2011, 1018): citizens should have the opportunity to learn about many different viewpoints on various issues, so that they can form a well-informed and thoughtful opinion of their own, and vote accordingly.
Following this ideal within an electoral democracy, we focus on the coverage of the competing political parties. News that is perfectly diverse provides citizens with information about all political parties - their viewpoints, their successes or failures of candidates to an equal extent during an election campaign (open diversity, see Van der Wurff and Van Cuilenburg 2001). Differences between outlets can reveal the extent to which the outlet also pays attention to the views of smaller parties. The more an outlet focuses on some political parties over others, the greater the chance for some kind of bias, and the lower the diversity.

For online newspapers, more space could further diversity. So far, however, only Powers and Benson (2014) have demonstrated that this is the case. Moreover, the interactive elements of a website could increase diversity through facts or opinions supplied by the readers themselves in the forum or comments section below news stories in online editions. This input by the audience may then influence news selection and presentation by journalists. Some researchers point out that this may be particularly true for tabloid newspapers because it is their aim to echo the voice of the people rather than the establishment, unlike elite papers (e.g. Örnebring and Jonsson 2004; but doubts in Örnebring 2006 and Thurman 2014). For the diversity of party coverage, which we measure here, this could mean a lower focus on government parties and more attention to the opposition, including new and populist parties, resulting in higher diversity. The interactive features of online media should make it even easier for popular newspapers to function as an alternative public sphere, as they make it easier for outlets to learn about people’s viewpoints. However, online media appear not to have lived up to this potential (Örnebring 2008; Conboy and Steel 2010; Richardson and Stanyer 2011).

This could be because of the simultaneous process of commercialization. As we argued earlier, audience metrics make it easier for outlets to cater to audience preferences, also with respect to specific political preferences (Tandoc 2014). In addition, a faster news cycle means that journalists have less time to reflect on what they publish and which angles of a story to choose. So, biases of news
selection based on routine, i.e., news values ingrained in journalistic training, are more likely to surface (Welbers et al. *in press*). This leads to the following research questions:

**RQ2a** *How high is the diversity of political news in online and in print newspapers?*

**RQ2b** *Is there a gap in the degree of diversity of political news in popular and in elite newspapers, and is the size of this gap similar online and in print?*

*Focus on Leaders*

A third indicator considers the focus on political leaders, or presidentialism (Langer 2007), a type of personalization of political news (Van Aelst et al. 2012). Political leaders, being powerful elite persons, are high in news value (Galtung and Ruge 1965; Harcup and O’Neill 2001). So, they are relatively likely to be written about in the media, at the cost of politicians lower in status or of political institutions. However, from a democratic perspective a high leader focus is less desirable. It leads to personalized voting – voters who weigh evaluations of political leaders more heavily than issue positions (Takens et al. 2015) which goes against the ideals of the electoral democratic model, which holds both as important (Strömbäck 2005). Also, a stronger focus on leaders means less attention to other candidates of the same party, whom people may also vote for in countries with a preferential vote system such as Austria. One could thus argue that a strong focus on leaders goes against the political logic of such systems (Rahat and Sheafer 2007, 66), and regard a low focus on leaders as a sign of news quality by electoral democratic standards.

Commercialization of news is seen as an important explanation for the personalization of political news (Langer 2007). Personalizing and simplifying news are ways to make it more attractive to a large audience, and focusing on leaders does both (Dahlgren 2000, see also Bird 2000 and Eilders 2006 for audience perspectives). A more competitive media environment in which news outlets follow these
preferences more closely, such as the one found online, may therefore lead to a greater emphasis on leaders in times of election campaigns (see also Takens et al. 2015).

**H1** *Political news is more focused on leaders online than in print.*

Popular newspapers, as mentioned previously, cater to the tastes of an audience that is as large as possible. This makes it likely that focus on leaders is stronger in popular newspapers than in elite newspapers, which presumably focus on making ‘quality’ news, for an audience that supposedly prefers ‘quality’ news. Whether this holds true to a similar extent online and in print depends on how strongly elite newspapers are affected by the commercialization pressures of their online version.

**RQ3** *Is there a gap between the strength of leader focus in the political news in popular and in elite newspapers, and is the size of this gap similar online and in print?*

**Emotionalization**

In order to sell news to audience and advertisers, journalists may focus on those aspects of news events that people can identify with. This may also result in a greater use of emotions in their news coverage (Grabe et al. 2001; Donsbach and Büttner 2005). This, however, goes against the detached style of political news coverage preferred by the electoral democratic model (Jandura and Friedrich 2014), which evaluates political coverage low in emotionalization as higher in quality.

As with leader focus, more competition online and more possibilities to directly track audience preferences could result in more emotionalization in online media coverage.

**H2** *Political news is more emotionalized online than in print.*

Again, popular newspapers are focused on following the tastes of a large, broad audience, with different preferences than the elite audience for elite media, which likely results in more emotionalized coverage in popular media (Reinemann et al. 2012). Whether this
holds true to a similar extent online and in print editions again depends on how strongly elite newspapers are affected by the commercialization possibilities of their online version.

**RQ4** *Is there a gap between the degree of emotionalization of political news in popular and in elite newspapers, and is the size of this gap similar online and in print?*

**Case Study, Data, and Methods**

This study analyzes the full political coverage of the online and print editions of six popular and elite newspapers, from August 19th until September 29th, 2013 (N=14,868), during the Austrian National Election Campaign for that year². None of these papers uses paywalls; all online content is freely available for everyone. Except for the elite paper *Der Standard*, all newspapers also appear on Sundays. Since we focus on news supply during the overall campaign, having a Sunday edition or not does not matter for our comparison, meaning that *Der Standard* has one paper per week less than the other outlets in our sample.

‘Political coverage’ includes all news stories that mention the election, an Austrian politician, political party or other political institution. The print newspaper coverage was downloaded from the Austria Presse Agentur (APA) database,³ the online coverage was scraped daily, from the newspaper websites for the previous day. A news story that is updated or changed later on the same day is thus included as its final version, and counted only as one news story. However, if it is updated or edited again on the next day, both versions are counted.

All four indicators for democratic news quality are measured using automatic content analysis (search strings) in an online content analysis toolkit (http://amcat.nl; see Van Atteveldt 2008), providing

---

² To be published in the GESIS data archive (http://www.gesis.org/) by the end of 2015.
³ http://www.apa.at/Site/index.de.html
high reliability. Using automatic rather than manual content analysis makes it possible to include a larger amount of news coverage.

In addition to the quantitative content analysis described above, we also compared the general format of each print newspaper to its online edition to see which sections (such as letters to the editor, or opinion columns) overlapped or were missing from either version in order to contextualize our quantitative findings.

**Operationalization**

In the following section, we will elaborate on the indicators used to map the normative criteria discussed above. For the **amount of political news**, we used both the number of news stories on politics published by a particular outlet as well as the average length (in number of words) of these stories. **Diversity** was measured by automatically coding whether a party is mentioned within a news story or not. For this, we used search strings, that is, words or combinations of words that signify a specific party, so, e.g. SPÖ, but also social democrats. We only included the nine political parties that took part in the national elections in all Austrian states (Bundesländer). The average precision for these search strings is 0.98, with no value lower than 0.88, and the average recall 0.95, with 0.86 as its lowest value.

The measure we used to calculate diversity is entropy in number equivalents, or the perplexity transformation of Shannon’s H (Kleinnijenhuis et al. 2015), which was calculated using the following formula:

\[
diversity(n) = \prod_{i=party}^{n} \left( \frac{1}{p_i} \right)^{p_i}
\]

where \( p_i \) stands for the proportion of media attention for a particular party relative to the media attention for all parties (Van Hoof et al. 2014). This transformation results in an easy-to-interpret score with

\(^4\) For the exact search strings, see Haselmayer et al. (forthcoming).
a minimum of 1 (all media attention goes to one party) and a maximum that is equal to the number of categories for that variable, in our case, $n = 9$ (all nine parties get exactly the same share of media attention).

For **focus on leaders** we used search strings to automatically code whether a political party and/or a top candidate of a political party was present in a news story. A top candidate is the number one on the candidate list, who is often (but not by definition) also the party leader. Average precision for the search strings is 0.99 with 0.88 as the lowest score, and average recall is 0.95, with 0.64 as the lowest score.\(^5\) ‘Leader focus’ was then operationalized as the share of news stories mentioning a top candidate out of all news stories that mention a political party or its politician(s). As for diversity, we included only the nine political parties that took part in the national elections in all Austrian states.

**Emotionalization** was measured as the share of emotion words out of all words used in the coverage by a particular outlet (see Cho et al. 2003). We created a search string consisting of a slightly modified version of the sentiment lexicon Sentilex (Wolf et al. 2008) to measure the number of sentiment or emotion words (words referring to or provoking emotions) occurring in a news story. Sentilex consists of all positive and negative sentiment words – nouns, adjectives and verbs—from a larger dictionary used in automatic text analysis, mostly in the domain of psychology, the German-language version of Linguistic Inquiry and Word Count (LIWC, Pennebaker et al. 2001; Wolf et al. 2008). We modified the entries in this dictionary slightly in order to use it as a search string, using wildcards (e.g. ‘verlier*’, ‘to lose’ or ‘loser’) wherever possible in order to include composite

\(^5\) This one low score is for Frank Stronach, leader of the political party Team Stronach, where the similarity between the party name and the leader name leads to difficulties in distinguishing between the two using search strings. As no easy solution for this problem is available, and precision and recall values for all other search strings used for the leader focus variable were very high, we chose to keep the search string and the analysis as it is.
words and conjugations. Where this was not possible (e.g. searching for ‘frei*’, ‘free*’ in our data results in a large number of hits for ‘Freitag’, ‘Friday’, which is not a sentiment word), we either entered all grammatical conjugations and relevant words separately, or chose to delete the word from the dictionary. The precision for this search string was 0.90, and the recall 0.89. We then calculated the degree of emotionalization, that is, the percentage of sentiment words in the political coverage (measured in words) of a particular outlet.

Since we used the full political coverage as opposed to a sample, we conducted no significance tests.

**Differences between Print and Online Editions**

Our first research question asked how political news during an election campaign in online editions in general compares to that in print editions in general. The results of this comparison using our full media sample are shown in Table 3.1.

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Print</th>
<th>Online</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average number of political stories per outlet</td>
<td>1,587</td>
<td>891</td>
<td>-697</td>
</tr>
<tr>
<td>Average story length in words</td>
<td>237</td>
<td>448</td>
<td>+211</td>
</tr>
<tr>
<td>Party diversity (Entropy $E$, scale: 1-9)</td>
<td>6.27</td>
<td>6.74</td>
<td>+0.47</td>
</tr>
<tr>
<td>Stories focusing on a party leader (%)</td>
<td>35.6</td>
<td>47.2</td>
<td>+11.7</td>
</tr>
<tr>
<td>Degree of emotionalization</td>
<td>4.5</td>
<td>4.0</td>
<td>-0.5</td>
</tr>
</tbody>
</table>

Online newspapers contain fewer news stories about politics. Their news stories are, however, almost twice as long, meaning that altogether, online editions provide more political coverage than print editions. If we take a closer look at these news stories, we see that online media tend to publish very long (over 2,000 words) political stories regularly. They are often live blogs, stories that consist of photo slideshows with text underneath, or stories that reproduce tweets of politicians in addition to a core text. Letters to the editor were missing from the online edition.
Diversity is quite high online, with a score of $E = 6.74$ out of 9. In print media it is a little bit lower at $E = 6.27$. Online editions pay slightly more attention to opposition parties and new parties than print editions. A focus on party leaders, however, is also more frequent online than in print outlets, confirming our hypothesis regarding this indicator ($H1$).

Lastly, emotion words are used slightly less often online than in print. Our hypothesis regarding this indicator is thus not confirmed ($H2$). Overall, the picture emerges of a different kind of news coverage online as compared to print, with fewer but longer news stories, in which more parties are included but that also focus more strongly on the leaders of these parties, and that contain fewer emotion words.

**A Gap in Quality between Popular and Elite News Media?**

In the next section, we investigate the potential gap in quality between popular and elite newspapers in their print and online editions.

*Amount of political news*

Table 3.2 shows our findings regarding the amount of political news stories during the election period, as well as their average length.

Somewhat surprisingly given their image of being low in political news, popular newspapers have more political news stories than elite newspapers in print, about twice as much in total. However, online this gap is reversed, and strongly so: here, elite newspapers contain three times more political stories than popular newspapers.

Popular newspapers do publish longer news stories online on average, although this average is strongly influenced by one outlet, and does not compensate for the vastly lower amount of news stories, both compared to elite newspapers as well as to popular newspapers’ own print editions.
Table 3.2 Amount of political news in popular and elite newspapers, print and online editions

<table>
<thead>
<tr>
<th></th>
<th>Number of Stories</th>
<th>Average length</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Print</td>
<td>Online</td>
</tr>
<tr>
<td>Elite</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Der Standard</td>
<td>998</td>
<td>1493</td>
</tr>
<tr>
<td>Die Presse</td>
<td>1191</td>
<td>1159</td>
</tr>
<tr>
<td>Salzburger Nachrichten</td>
<td>834</td>
<td>1372</td>
</tr>
<tr>
<td>Average</td>
<td>1008</td>
<td>1341</td>
</tr>
<tr>
<td>Popular</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kronen Zeitung</td>
<td>2253</td>
<td>269</td>
</tr>
<tr>
<td>Kurier</td>
<td>1694</td>
<td>638</td>
</tr>
<tr>
<td>Österreich</td>
<td>2554</td>
<td>413</td>
</tr>
<tr>
<td>Average</td>
<td>2167</td>
<td>440</td>
</tr>
<tr>
<td>Difference Popular — Elite</td>
<td>+1159</td>
<td>-901</td>
</tr>
</tbody>
</table>

In sum, differences between media types in the amount of news are larger online than they are in print, to the advantage of elite newspapers.

Aside from this change in quantity, the types of news stories included in especially popular newspapers are also different online. In two of the three popular newspapers, certain sections prominent in the print edition were not included in the online editions, namely the opinion and letters from the editor sections. The other four (three elite and one popular) papers featured a discussion forum (not counted as news stories) instead of a section for letters to the editor, but these papers did publish opinion columns in their online edition.

Diversity

For diversity, we saw that online versions in general are slightly more diverse than print versions. Table 3.3 gives an overview of party diversity in popular and elite newspapers, measuring how evenly media attention is distributed over the nine political parties that took part in the 2013 election.

Nearly all individual outlets are more diverse online than in print. In the print editions, the elite newspapers tend to show more variation
than the popular newspapers. Online, however, popular outlets catch up and are equally, if not more, diverse than elite outlets.

For this indicator, differences between media types in their diversity are smaller online than in print. Popular newspapers even slightly surpass elite newspapers in their online edition. Elite media provide more political news online than popular media, but all parties do not benefit equally: they provide many more stories about some parties, rather than a little bit more coverage of all.

**Table 3.3** Party diversity in popular and elite newspapers, print and online editions

<table>
<thead>
<tr>
<th></th>
<th>Print</th>
<th>Online</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Entropy</td>
</tr>
<tr>
<td>Elite</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Der Standard</td>
<td>808</td>
<td>6.65</td>
</tr>
<tr>
<td>Die Presse</td>
<td>877</td>
<td>6.50</td>
</tr>
<tr>
<td>Salzburger Nachrichten</td>
<td>697</td>
<td>6.05</td>
</tr>
<tr>
<td>Average</td>
<td>6.40</td>
<td></td>
</tr>
<tr>
<td>Popular</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kronen Zeitung</td>
<td>1502</td>
<td>6.01</td>
</tr>
<tr>
<td>Kurier</td>
<td>1227</td>
<td>6.31</td>
</tr>
<tr>
<td>Österreich</td>
<td>1882</td>
<td>6.11</td>
</tr>
<tr>
<td>Average</td>
<td>6.14</td>
<td></td>
</tr>
<tr>
<td>Difference Popular—Elite</td>
<td>-0.26</td>
<td></td>
</tr>
</tbody>
</table>

*Note: N stands for the number of stories mentioning one or more political parties; diversity is measured in entropy in number equivalents (min. 1 – max. 9).*

**Focus on leaders**

The indicator ‘leader focus’ measures how strongly political coverage focuses on political leaders, at the expense of other politicians or parties. The results are shown in Table 3.4.

Leader focus is stronger online than in print for all outlets, with the exception of elite paper *Der Standard*. However, this difference is much larger for popular newspapers than for elite papers. In the print edition, popular and elite newspapers have a similarly strong focus on leaders.
Table 3.4  Leader focus in popular and elite media, print and online editions

<table>
<thead>
<tr>
<th></th>
<th>Print</th>
<th>Online</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Leader focus (%)</td>
<td>Leader focus (%)</td>
</tr>
<tr>
<td>Elite</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Der Standard</td>
<td>808</td>
<td>1293</td>
</tr>
<tr>
<td>Die Presse</td>
<td>877</td>
<td>1042</td>
</tr>
<tr>
<td>Salzburger Nachrichten</td>
<td>697</td>
<td>1195</td>
</tr>
<tr>
<td>Average</td>
<td>34.6</td>
<td>37.3</td>
</tr>
<tr>
<td>Popular</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kronen Zeitung</td>
<td>1502</td>
<td>205</td>
</tr>
<tr>
<td>Kurier</td>
<td>1227</td>
<td>563</td>
</tr>
<tr>
<td>Österreich</td>
<td>1882</td>
<td>380</td>
</tr>
<tr>
<td>Average</td>
<td>36.6</td>
<td>57.2</td>
</tr>
<tr>
<td>Difference Popular—Elite</td>
<td>+2.0</td>
<td>+19.9</td>
</tr>
</tbody>
</table>

Note: N stands for the number of news stories mentioning one or more political parties or its politician(s); leader focus is the share of these stories that mention one of more top candidate(s).

Online, however, this focus is strongly increased in popular newspapers. Notably, as much as 64.5 percent of all news stories about a political actor in *Kurier* mention a top candidate.

In both popular and elite newspapers, leader focus is stronger online, but the difference in leader focus is much greater for popular papers than for elite papers. As the total amount of news stories in popular media was also much smaller online than in print, it is more often the news about top candidates that makes it into the online edition, while the news about other political actors does not.

**Emotionalization**

The final indicator, emotionalization, is represented by the percentage of emotion words in the overall news story text in political coverage, shown in Table 3.5.

In every outlet, the online version is less emotionalized than the print version. In the print edition, emotionalization is similar in popular and elite papers. Online, even though the amount of coverage is very different, both popular and elite newspapers show less
emotionalization in their political coverage, a small difference of around 0.5 percentage points.

Furthermore, the emotionalization of news differs according to individual outlet rather than by type. Especially the *Kronen Zeitung* has less emotionalized political coverage online than in print, while for *Österreich*, online and print are emotionalized to an equal extent.

**Table 3.5** Emotionalization in popular and elite media, print and online editions

<table>
<thead>
<tr>
<th></th>
<th>Print</th>
<th>Online</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Average Length</td>
<td>Emotion words (%)</td>
</tr>
<tr>
<td><strong>Elite</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Der Standard</em></td>
<td>298</td>
<td>4.5</td>
</tr>
<tr>
<td><em>Die Presse</em></td>
<td>307</td>
<td>4.5</td>
</tr>
<tr>
<td><em>Salzburger Nachrichten</em></td>
<td>293</td>
<td>4.5</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td>300</td>
<td>4.5</td>
</tr>
<tr>
<td><strong>Popular</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Kronen Zeitung</em></td>
<td>175</td>
<td>5.0</td>
</tr>
<tr>
<td><em>Kurier</em></td>
<td>229</td>
<td>4.4</td>
</tr>
<tr>
<td><em>Österreich</em></td>
<td>121</td>
<td>4.3</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td>175</td>
<td>4.6</td>
</tr>
<tr>
<td><strong>Difference Popular - Elite</strong></td>
<td>+0.1</td>
<td></td>
</tr>
</tbody>
</table>

*Note: Average length shows the average word length of news stories in an outlet; emotionalization expresses the percentage of emotion words out of all words in the text of these stories.*

Differences in the emotionalization of political coverage in popular and elite newspapers are similar online and in print: for both, the share of emotion words in their coverage is about a half a percentage point less online than in print.

**Discussion and Conclusion**

According to the electoral model of democracy, news media should provide sufficient political information which is diverse, not too strongly focused on leaders, and that is detached rather than emotionalized.
The online media environment is both a challenge and an opportunity for political news by these standards. Among the challenges are the hasty news cycle and the increased competition compared to the print market. Virtually unlimited space and interactivity online, however, carry great potentials for news high in quality. In our study, we found that online newspapers mainly take advantage of these possibilities. They provide their audiences with fewer, but longer news stories, resulting in more political news overall. Online news also showed higher party diversity and less emotionalization. Rather than following a commercial logic, online newspapers in general provide a better coverage by electoral democratic standards.

However, by the same quality standards, the stronger focus on leaders in online editions makes online editions score worse than their print equivalents. An explanation for this focus could be the discussions of televised debates between top candidates in the media (Dolezal et al. 2014), which received even more attention in the online editions, as they could cover the debates in real-time. As leader focused coverage may lead to personalized voting rather than voting based on issue preferences (Takens et al. 2015), this could be a problem, especially if online coverage does not associate political leaders with their policy position (which could support issue-based voting).

Although our hypothesis regarding higher emotionalization online was not confirmed, an alternative explanation could be tested in further research. Namely, another assumed characteristic of online news is its high reliance on press agency material (Boczkowski and De Santos 2007), which is presumably low in emotionalization. News story types unique to online editions, such as live blogs and photo slideshows, also appear to be low in emotionalization, although we did not study this separately. Perhaps the different news sources and news story types unique to online coverage thus explain its lower emotionalization compared to print newspapers. Whether this is true or not, and what the implications of this would be for its quality, would be a good topic for follow-up research.
Grouping all print newspapers or all online newspapers together, however, obscures differences between specific outlets in their approach to publishing online versions, as indicated by research in the United Kingdom. Taking a closer look at the individual news outlets, we found that the print version of popular newspapers contained more political news than the print version of elite newspapers, although their diversity was slightly lower. This is consistent with the finding of Magin and Stark (2014) who noted the high political involvement of the ‘people’s newspaper’ Kronen Zeitung. Our data suggest that this characteristic may apply to other Austrian popular newspapers as well, but only in their print edition. Leader focus and emotionalization were comparable to that in elite newspapers. In print, political news in elite and in popular newspapers is therefore rather similar in quality as measured by our indicators, although popular newspapers do even slightly better given their greater amount of political news.

The performance of Austrian news outlets in online editions seems to be in line with studies of tabloids and tabloidization (e.g. Sparks 2000): popular news outlets provide less political coverage than elite outlets. One reason for the lower amount of political news stories in online popular newspapers could be that especially for these papers, the online versions do not include certain sections that are prominent in the print edition.

Furthermore, we see this as a sign of a stronger influence of commercial logic in online editions of popular newspapers which focus strongly on those politicians that everyone knows and will recognize, instead of also giving other candidates a larger share of attention. Elite newspapers show this sign of commercialization less strongly.

The slightly higher diversity in popular newspapers is, however, not what we would expect following a commercialization scenario, but rather seems to confirm the view of popular media as paying more attention to alternatives than to the status quo (Örnebring and
Jönsson 2004). However, this holds only true for online editions, and only slightly so.

Lastly, the degree of emotionalization is fairly equal in elite and popular newspapers online, again corresponding to Magin and Stark’s (2014) findings about Austrian newspapers. Rather than commercial logic, an explanation for the lower amount of emotionalization online, as mentioned previously, could be the live blogs, tweets and press agency material – descriptive or fact-oriented news stories rather than the opinion pieces, analyses and reader letters, that are important in the print version. The lower use of emotion words hence appears to be a general characteristic of online news. Thus even though elite and popular newspapers show a number of similarities in their adaption of online news formats (greater diversity, lower emotionalization), elite newspapers are more successful than popular newspapers in providing good quality in their political coverage online: they offer much more of it, and focus less on party leaders.

Overall, differences between popular and elite newspapers tend to be larger online than in print, to the disadvantage of popular newspapers. As such, our data seem to confirm what Sparks (2003; see also Rucht, Yang, and Zimmermann 2008) found: the move from print towards online media may reinforce the gap between a well-informed elite audience and a lesser-informed audience for popular news, especially since more young people prefer online news media over other channels (e.g. Mitchelstein and Boczkowski 2010). By corroborating the results for the United Kingdom for a country with a different political communication culture – more partisan news outlets (Lengauer and Johann 2013; Eberl, Boomgaarden, and Wagner 2015) and a different political system with proportional representation and a large number of relevant parties – our Austrian case study encourages us to assume that these findings can be generalized to a larger context. However, further research is needed, particularly comparative research, to explore possible explanations for the gap in quality between political news in popular and elite papers.
Furthermore, within this article, we focused on a specific democratic model and a corresponding set of indicators that could be assessed via automatic content analysis. This, however, represents only a limited picture of what can be defined as ‘political news quality’, and only for the specific context of providing people with political information during election campaigns. We encourage future projects to explore the contribution of print and online media to democracy following other understandings of the role of media in a democracy so that a more complete picture may emerge over time, when possible combining automatic and manual content analysis in order to benefit from the advantages of both respective approaches.

Other interesting findings in our study suggest that the orientation function – providing readers with commentary and explanation rather than news facts only (Connell 1998) – of especially popular newspapers is at peril online, since opinion sections are not always included and analysis and commentary are de-emphasized. We argue that this function cannot be taken over by the additional sections unique to online versions, such as live blogs, photo collections, or reader fora, as in particular the latter may provide lots of opinion, but little structure and orientation for its readers. While online editions in general thus perform slightly better than their print counterparts following the quality criteria used here, they have their own specific pitfalls, which need to be examined further in order to understand their (potential) consequences for democracy.

References


The Quality of Political News in a Changing Media Environment


3. Political News Content in Online and Print Newspapers


The Quality of Political News in a Changing Media Environment


3. Political News Content in Online and Print Newspapers


The Quality of Political News in a Changing Media Environment


Tandoc, Edson C. (2014). Journalism is twerking? How web analytics is changing the process of gatekeeping. *New Media & Society*, 16(4), 559-575. Doi: 10.1177/1461444814530541


3. Political News Content in Online and Print Newspapers


Welbers, Kasper, van Atteveldt, Wouter, Kleinnijenhuis, Jan, Ruigrok, Nel, and Schaper, Joep (in press). News selection in the digital era: professional norms vs. real-time audience metrics. Accepted for publication in Journalism.

Abstract

In this chapter, we explore how network analysis can be used to analyze the complexity of the content of political coverage in newspapers. By conceptualizing political coverage as a mediated network representing a public sphere, we are able to adapt three indicators from network analysis to compare and evaluate the content complexity of political news in different outlets. Network density enables us to judge the general interconnectedness of the depicted debate; network centralization allows us to look at the focus on specific actors or issues; and network heterogeneity reflects the diversity within these actor and issue networks. Following a theoretical and methodological discussion of the applicability of these three indicators for evaluating news complexity, we demonstrate their usefulness in a case study comparing the political news coverage of tabloid and elite newspapers during a national election campaign.

6 A modified version of this article is in review for publication in the journal European Journal of Communication.
Introduction

News complexity has traditionally been regarded as a quality criterion for news media: ‘quality’ outlets supposedly have complex coverage, whereas low-quality tabloids are generally assumed to simplify their news, and thus to be less complex than elite newspapers (Connell 1998). In empirical investigations, news complexity is mainly reduced to word difficulty or syntactical complexity, without taking into consideration the actual meaning conveyed by these words and sentences, that is, the picture of the world that they depict. To fill this gap, this article proposes complexity measures which can assess the complexity of news content. By conceptualizing political coverage as a mediated network representing a public sphere, we can draw on indicators from network analysis to evaluate the complexity of this network and thus news content complexity.

Network analysis, both social and semantic, has been used before in communication research (e.g. Kleinnijenhuis and de Nooy 2013; De Nooy and Kleinnijenhuis 2013; Takens et al. 2013; Kleinnijenhuis et al. 2015), but its measures have thus far not been applied to investigate the complexity of news content. This is the case even though the measurement of the complexity of networks represents a central question in network analysis (specifically, graph theory), which has hence developed various complexity measures. We argue that these methods can also be fruitfully applied to the complexity of political news coverage, which we see as a mediated network (similar to “symbolic network” as used in Adam 2008) of the public sphere as represented in a media outlet. We use the term ‘mediated network’ to emphasize that these networks occur in media reporting, as opposed to ‘real world’ networks between politicians or other actors who communicate with each other (ibid.). As such, these mediated networks serve as the site where we measure this indicator of news quality in the respective media outlets – content complexity.

In the following sections, we will first discuss the normative assumptions on which complexity as a criterion for news coverage is
based, as well as previous attempts of evaluating news complexity. We then develop a definition of content complexity and propose to measure this form of complexity using three indicators drawn from network analysis. To illustrate how network measures can be used for measuring news content complexity, we will apply our measures in a case study comparing the political news coverage of Austrian popular and elite newspapers during Austria’s 2013 national election campaign, based on a data set of manually coded political statements (N = 2,763). In the final section, we will outline additional applications outside the political news realm and promising avenues for future research.

**Complexity as a Quality Criterion for News Coverage**

Though complexity has often been put forward as quality criterion for news coverage, we still lack a concise definition which integrates both the underlying normative assumptions regarding news quality and the empirical dimensions of complexity in concrete news reports.

From a normative point of view, complexity of news coverage is deemed important based on the following three assumptions. First, politics is complex: new policy proposals or changes in existing laws may have many consequences, possibly also in policy areas beyond those of the original proposal (Jervis 1997). Politics is also complex as a process: consider the different opinions among parties on a particular policy proposal, including who agrees or disagrees with whom on which details, and how these change as the proposal develops from a first draft to a law. Second, citizens should be informed about politics in all its complexity. This is the ideal of an electoral democracy which requires citizens to be informed about politics (Dahl 1998; Strömbäck 2005). In practice, this ideal of fully informed citizens is not possible, as the resources of media outlets are limited, and so are the time and political interest of citizens.

Which leads us to the third assumption: real-world complexity has to be reduced in media reports. Though necessary, this reduction of complexity or rather simplification of news coverage can be seen as a problematic symptom of media logic, especially when it happens in all media, instead of only in non-elite outlets. The media logic thesis
states that (political) media coverage is heavily influenced by news values and format requirements (Mazzoleni 1987; Takens et al. 2013). Simplification is one of these requirements. Journalists or news editors simplify events so that they fit better into a standard narrative or format (Altheide 2004). This also involves leaving out details that do not fit into the story that the journalist wants to tell. So, simplification can also be used as to sell more newspapers or attract more viewers, either deliberately or unconsciously as a result of norms and practices ingrained with training (Strömbäck and Kaid 2008).

From this normative perspective, complexity as a quality criterion thus addresses the question how well the media coverage ‘mirrors’ the ‘real world’ of politics, in other words it applies to the content, the actors, issues, as well as the relations and interconnections between them depicted in news coverage.

The reduction of news complexity, however, can also be seen as a way to enhance accessibility for audiences with a lower education, and by doing so, to encourage their interest in politics. In this view, the goal of journalism is not only to transmit information from political processes to citizens, but also to make sure that this information is well understood by their audience (Dahlgren 2000). This involves for instance presenting only the main point of the news and leaving out details (i.e. peripheral actors or related policy issues) that may render the understanding of the main message more difficult for particular audiences (Gans 2009) – in other words, to make news less complex.

By defining complexity through its possible consequence, (lack of) understandability, this perspective draws attention to the lexical or syntactical level of news coverage: Easy sentences and words are assumed to enhance people’s ability to understand what is being said on the news, complex sentences to hamper it (Blom 2011).

For this reason, many empirical investigations of news complexity focus on these (easier to operationalize) concrete textual elements: They rely on word difficulty as an indicator of news complexity (e.g. Donsbach and Büttner 2005) or measure the length of sentences or of sound bites (e.g. Hallin 1992), assuming that longer sentences are
grammatically more complex than short sentences. Blom (2011) analyzed the syntactical complexity of radio news by mapping sentences as a kind of network: syntactical complexity is high when within sentences referents are used that are far apart from the words they refer to (ibid.). However, these measures are based either on single words, or on structural or grammatical aspects of sentences or texts. None of them measure the complexity of the content depicted in these texts directly. By contrast, it is the aim of this article to develop a definition and indicators for content complexity, which is at the heart of normative demands for news complexity discussed above.

**Defining Content Complexity**

The question whether the content of media coverage shows the complexity of political reality in an appropriate manner has been addressed before mostly with respect to diversity. Diversity concerns itself with the variety and the balance of attention distribution between stakeholders and issue categories in the news, and the more evenly news attention is distributed between these categories, the higher diversity (McDonald and Dimmick 2003). For example, Kleinnijenhuis et al. (2015) analyze the development of the diversity of issues and frames in the news coverage over time as agenda and frame complexity: The more diverse the media content, that is the more issues and frames are discussed, the higher agenda- or frame complexity. Similarly, Kleinnijenhuis (1991) analyzes complexity as the extent to which yesterday’s news agenda can predict that of today.

However, diversity studies only consider the occurrence of actors, issues or frames, irrespective of their context, that is, the interconnections between them. It is our opinion that content complexity should address the connections that are explicitly drawn between actors and issues. After all, politics is all about the relationship between different actors and issues: who promotes a specific policy goal? And who opposes this initiative by pointing out its consequences in another issue area? An evaluation of news content complexity which limits itself to the analysis of occurrences would thus not do justice to the complexity of politics.
Table 4.1 Overview of complexity indicators

<table>
<thead>
<tr>
<th>Level of analysis</th>
<th>Object of analysis</th>
<th>Indicators</th>
</tr>
</thead>
<tbody>
<tr>
<td>Textual complexity</td>
<td>Word</td>
<td>Word difficulty (e.g. Donsbach and Büttner 2005)</td>
</tr>
<tr>
<td></td>
<td>Sentence</td>
<td>Length of sentences/soundbites (e.g. Hallin 1992)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Distance of referents in sentences (e.g. Blom 2011)</td>
</tr>
<tr>
<td>Content complexity</td>
<td>Occurrences</td>
<td>Diversity (e.g. Kleinnijenhuis 1991)</td>
</tr>
<tr>
<td></td>
<td>Network</td>
<td>Interconnectedness</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Centralization</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Heterogeneity</td>
</tr>
<tr>
<td></td>
<td>Argument structure</td>
<td>Integrative complexity (differentiation and integration, e.g. Suedfeld and Tetlock 1977)</td>
</tr>
</tbody>
</table>

Even more importantly, the organization of knowledge in texts has been shown to have an impact on cognitive processing. According to Graber (2001), people can obtain both denotative information, the simple recording of facts, and connotative information from news texts. The latter refers to the assignment of meaning to the observed actors and issues, which is based on the connections between the different elements of information in the text. From a normative perspective, this connotative information is the more valuable, as it allows people to reason effectively and reach political decisions reflecting their own interests.

Eveland, Marton and Seo (2004) provide initial evidence that a higher ‘knowledge structure density’, that is, the greater “number of interconnections across concepts” (Eveland, Marton and Seo 2004, 88), in online news texts with hyperlinks increases the density of the interconnected knowledge structure of their readers, that is their connotative knowledge. Hence, it should be the more complex network of information provided in news texts from which readers
should be able to draw more connotative information. And this in turn should improve their ability to reason.

This is why we investigate the complexity of mediated networks of political claims – that is, statements by actors about other actors or issues within political news (Koopmans and Statham 1999). Here, we look beyond single sentences: actors may be connected to actors in another sentence when the context makes clear that this is whom they are talking about. For connecting issues, we look at the entire news story. Through this we abstract from single sentences and analyze the structure of the public discourse represented in a particular outlet as reflected in the relations between actors or issues (see also Adam 2008).

With regard to content complexity as a quality criterion of news coverage, we define it as the complexity of the mediated network of interconnections of actors and issues as depicted by a media text. From a normative perspective that expects media to mirror politics as closely as possible, high content complexity (or a low level of simplification) would be seen as advantageous, because the reality of politics is complex. A highly complex media representation of this reality is likely to be a good representation of political debates and may encourage the acquisition of connotative political knowledge. There is no similar argument in favour of high syntactical complexity: arguably, complex debates can also be described in simple sentence structures, and thus made easier to understand for a wide audience.

---

7 To our knowledge, the only other complexity measure which relates to both the content and its structure, is the integrative complexity measure used in political psychology (see e.g. Suedfeld and Tetlock 1977, Conway et al. 2014) which looks at the intellectual style of individual or groups: how many issue dimensions does one actor (or group) address and how strongly does he (or they) connect them? However, the assessment of the two main indicators of integrative complexity, differentiation and integration, is based solely on scorings by human coders for which reliability can only be achieved through rigorous training (see Suedfeld, Tetlock, & Streufert, 1992) – this makes the measure impractical for analyzing larger samples of (news) texts. And the recent attempt by Conway et al. (2014) to code integrative complexity (semi-)automatically is again based on words and sentence structure, in other words on textual complexity.
Indicators of Content Complexity Derived From Graph Theory

Understanding the content of news coverage as mediated networks allows us to consult graph theory and the more traditional applications of network analysis for indicators of (network) complexity that can fruitfully be transferred to the particular case of mediated networks and the measurement of news content complexity.

Interconnectedness

A first complexity measure commonly applied in network analysis is interconnectedness, that is, the number of connections between the different nodes of a network. In graph theory, both a graph in which no nodes are connected and a fully connected graph are considered to be low in complexity (Gell-Mann 1994, 31; Kim and Wilhelm 2008, 2638). The reason for this is that branching, the number of relevant substructures, and the heterogeneity of nodes, as well as other dimensions of graph complexity, would be very low in such graphs. In the mediated networks of political news content that we are concerned with, however, this is not the case. Here, the complete interconnectedness of all actors would be an indicator for a political public sphere where every actor is interacting with every other actor, or where every issue would have been discussed in the context of every other issue. For media content, this would be a highly complex situation, even though graph theory assumes otherwise (see also Bonchev and Buck 2005).

For the purpose of evaluating news content, we therefore propose to understand any increase in interconnectedness as an increase in content complexity. In actor networks, a strongly connected network would suggest that many actors are present in various debates either as speakers or addressees at some point. A complex network would be densely connected and have only few peripheral actors. In issue networks, a high interconnectedness of issues would imply that no set of issues is discussed in isolation from others. Policy issues may have diverse impacts on other policy fields. Communicating this in news coverage and thus raising several issues in the same news story will
result in a higher level of interconnectedness and a more complex representation of the political public sphere.

**Figure 4.1** Interconnectedness

In Figure 4.1, the nodes represent actors. In network 1, no actor is interacting with any other actor – in news coverage, this would mean that news stories discuss only one actor at a time, or depict only actors who do not engage in debate with one another. The resulting network would not be complex. Such a network is unlikely to occur empirically when mapping mediated networks of political news, as news item rarely fail to connect at least two actors.

In the second network, every actor is only connected to two of his neighbours, but no single actor is connected to more than two other actors. Actors $a$ and $b$ are engaged in a political debate with each other, and so are actors $b$ and $c$, and so on. In network 2, the number of debates each actor takes part in remains assessable. In network 3, however, every actor interacts with four different actors. We argue that it is much more complex to grasp each actor’s connection to each of her neighbours. This is also the case for the issue network, where a connection between node $a$ and $b$ signifies that the two issues were at some point being discussed in the same news story. Here, too, network 3 shows a population of more complex discussions than those presenting issues with only sparse connections between them or without any connections at all, represented by network 2 and 1, respectively.
Centralization

Another helpful indicator from graph theory is centralization which describes to what extent a graph is organized around particular focal points (Scott 2000, 89). In debates about news quality, ‘simplification’ – the opposite of complexity – is what is being feared. Simplification can simply mean a loss of political information, that is, parts of the political reality do not appear in the news at all. But simplification can also refer to news coverage that is excessively focused on only a small number of actors or issues. Even though lots of actors and issues may appear in such coverage, and they may have many connections, these connections are most often made between only a single or a few other actors and issues. In our understanding such a focus on a limited set of actors in the mediated network – in graph theory: a highly centralized network – would represent a simplification of political reality and thus a loss of complexity. By contrast, a debate where a multitude of actors/issues are discussed in relation to one another and where no clear focus is identifiable (low centralization), would be considered to be the more complex debate.

Figure 4.2 Centralization

In Figure 4.2, the first network is the most centralized. All nodes (actors/issues) have at some point been connected to the node in the middle, but not to any other nodes. By knowing about the storyline of the one central node and its connections, the structure of the whole network becomes clear. Arguably, this is a rather simplified representation of political reality, in particular if the network is based on more than one news item. The second network is a bit more
complex, as most actors/issues are connected to two other actors/issues, but some are only connected to one. As in network 1, media coverage is focusing on some nodes more than others. In the third network, the circle, all nodes are equally important in the structure of the network. To fully grasp this network’s structure, it is necessary to follow each node’s story line. This is the most complex and least centralized network presented here.

**Heterogeneity**

Lastly, in graph theory the heterogeneity of nodes in a network can also be seen as an indicator of network complexity (Kim and Wilhelm 2008). This should also apply to mediated networks: political coverage can be considered more complex if connections are drawn across different categories of actors and issues.

The way to use heterogeneity as a measurement of mediated network complexity employed here is by introducing network-exogenous factors (covariates). We want to see whether actors from different socio-political groups or issues from different political ideologies are connected to each other in the news. We argue that news coverage is more complex when actors from different social groups are shown to support or criticize one another, compared to news stories that focus solely on the view of a particular social group. Similarly, stories that discuss only issues belonging to a particular ideology (e.g. environmental protection, which is a ‘green’ issue) are less complex than stories that discuss issues from different ideologies together (e.g. news stories that discuss environmental protection in the context of tax reduction (an issue important in socio-economic rightist ideology), infrastructure (grey – the opposite of green ideology) and national pride (a nationalism issue)). The more heterogeneous the network, the more complex it is. This is illustrated in Figure 4.3, where the network on the right has more connections between nodes of different colours, and is thus more complex, than the network on the left.
Though heterogeneity is similar to diversity, diversity only looks at the occurrences of actors and issues, not at their interconnections: News coverage can be considered diverse if it contains statements from actors from different parts of the ideological spectrum. Heterogeneity as a dimension of content complexity, by contrast, is more demanding: here the actors would actually need to refer to each other.

Based on these three indicators, we can now expand our original definition of content complexity: Content complexity is the complexity of the mediated network of interconnections of actors and issues as depicted by the media text.

The mediated network in a particular outlet is more complex,

- the more connections are made in this outlet between actors or issues (interconnectedness)
- the less the mediated network is focused on a limited amount of actors and issues (de-centralization)
- the more actors (or issues) from different groups are connected (heterogeneity).

In a next step, we will apply our three indicators to an empirical case study. Here, we will demonstrate the usefulness of the different indicators for analyzing content complexity as the complexity of mediated networks of actors or issues, which allows us to say something about this aspect of political news quality in different outlets. We see our three indicators as separate, but not entirely
independent dimensions of complexity. A mediated network with few connections is unlikely to be very heterogeneous but likely to be more centralized. However, the relationship between the three dimensions is not likely to be linear either: an increase in connections between actors and issues would not automatically translate into an increase in heterogeneity or a decrease in centralization. Our empirical case study will allow us to probe the relationships between these different content complexity dimensions, and show why using all three (rather than just one indicator for complexity) is useful for our particular research goal.

Drawing on a manually coded dataset of political claims in national election coverage, we will compare the results of our three indicators for popular and elite news outlets. Political coverage of popular newspapers is often presumed to be less complex than that of elite newspapers, but it is unclear whether these assumptions are based on textual complexity or content complexity, and whether they are in fact based on current empirical reality. By employing our content complexity measures we will attempt to impartially map the degree of complexity in the mediated networks of the political news coverage in popular as well as elite newspapers.

Our research question for this case study reads:

**RQ** Is the content complexity of political news coverage in elite newspapers higher than that of political news coverage in popular newspapers, using our three dimensions interconnectedness, centralization, and heterogeneity?

**Data and Methods**

The data this study is based on was generated through a standardized content analysis using the Amsterdam Content Analysis Toolkit (AmCAT®), an open source platform that provides a framework for large scale automatic and manual content analysis. We analyzed the full political coverage in four Austrian national newspapers (two

---

tabloids and two elite papers) during the six weeks of the recent Austrian National Election campaign (August 19th through September 29th, 2013), a total of 2,763 news stories. We define political coverage here as all news stories in which at least one Austrian political institution, party or other actor is mentioned.

The content analysis method used in this study is a form of Semantic Network Analysis (Krippendorff 2013, 248) that combines elements of the NET-method for (Van Atteveldt 2008) and political claims analysis (Koopmans and Statham 1999). The units of analysis in this study were political claims. A claim is a statement by a speaker about a political issue and/or an addressee, through which he expresses critique, responsibility or approval (Adam 2008; Helbing and Tresch 2011). Two types of claims were coded: 1. Claims where an actor made a political statement or performed an action towards a political issue; i.e. “We (speaker) want better healthcare (issue)”. 2. Political statements or actions (in the context of policy) from one actor to or about another actor; i.e. “The social-democrats (speaker) do not want to collaborate with the FPÖ (addressee) on the issue of integration”. Inter-coder reliability for subject actors, object actors and issues was measured by having a subsample of the news stories coded by all coders (n=1,123 claims). Krippendorff’s α values for speakers, addressees and issues were 0.80, 0.72 and 0.74, respectively, which are acceptable values (Krippendorff 2013).

The actor and issue networks are constructed in different ways. Actors are connected with one another when speaking to or about each other in the same political claim. This is a one-mode network, where the nodes are referring to actors, and every political claim between two actors forms a tie between two nodes. A tie can thus represent an actor holding another actor politically accountable, agreeing or disagreeing with the other actor, or taking action against the other actor. As issues cannot be speakers, issues were connected to one another when they were mentioned in actor-issue claims within the same news story. This second network can be best understood as a co-occurrence network (a transformed two-mode network).
Because of this way of constructing networks, the direction of the links in the networks could not be taken into account for all our analyses. The co-occurrence network of issues, for example, is itself by definition undirected. Moreover, not all measures include the frequency of connections (edge weights). If not stated otherwise, links in these networks are binary.

For our purposes – comparing complexity in popular and elite newspapers – we analyze all networks on the level of individual news outlets, reflecting the public sphere as represented by that outlet over the course of the election campaign.

Table 4.2 shows an overview of the data used in this study. The elite papers analyzed in this study were Der Standard and Die Presse with 550 and 666 stories, respectively. The popular papers are tabloid Kronen Zeitung with the highest (n=1,193) and free paper Heute (n=354) with the lowest amount of political coverage. A total of 3,837 actor-claims and 15,262 issue-claims were coded, with an almost equal amount of claims in elite and popular papers.

Table 4.2 Data overview

<table>
<thead>
<tr>
<th>Outlet</th>
<th>News stories</th>
<th>Actor Claims</th>
<th>Issue Claims</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Percent</td>
<td>N</td>
</tr>
<tr>
<td>Der Standard</td>
<td>550</td>
<td>20.0</td>
<td>872</td>
</tr>
<tr>
<td>Die Presse</td>
<td>666</td>
<td>24.0</td>
<td>1,033</td>
</tr>
<tr>
<td>Kronen Zeitung</td>
<td>1,193</td>
<td>43.0</td>
<td>1,448</td>
</tr>
<tr>
<td>Heute</td>
<td>354</td>
<td>13.0</td>
<td>484</td>
</tr>
<tr>
<td>Total</td>
<td>2,763</td>
<td>100.0</td>
<td>3,837</td>
</tr>
</tbody>
</table>

Operationalization

To assess the previously discussed three dimensions for the complexity of actor and issue networks, we use the measures discussed below.

**Interconnectedness** A common concept in network analysis for measuring the interconnectedness of nodes in a network is density.
The density of a graph is defined by the number of links ($l$), expressed as a proportion of the maximum possible number of links, calculated based on the total number of nodes ($n$). The formula for density ($D$) is:

$$D = \frac{l}{n(n-1)/2}$$

The density of a network is 1 when all nodes are interconnected, and 0 when there are no connections between any nodes.

A problem with density is that it is highly dependent on the network's size (Scott 2000, 71ff). The more nodes are present in a graph, the lower the probability that all of them are connected to each other. Mayhew and Levinger (1976) argue that as the number of actors in a network grows, there is a saturation point. This is below the actual maximum number of possible links in a network (ibid.). For the issue network, too, one might say that it is unlikely that all issues affect each other and, therefore, are described as such in the media. Given this problem – larger networks have as a rule lower density than smaller networks – we chose to control for the size of the network. We will do so by calculating density not only for all actors and issues, but also for a sub-network based on the lowest common denominator (LCD) of actors or issues across all media. By taking into account only the actors/issues present in the networks of all media outlets, the actual number of nodes will be identical and therefore not influence the measurement of density.

**Centralization** Degree centralization measures the centrality distribution within a graph, investigating how focused a graph is on particular nodes. To calculate it, the centrality score of each node is subtracted from the node with the highest score. The formula for degree centralization is:

$$C(G) = \sum_n (\text{max}_w c_w - c_n)$$ where $c_n$ is the centrality of node $n$.

We normalized the result by dividing it with the theoretical maximum centralization, that is, the centralization scored by a star-shaped network (e.g. the first graph in Figure 4.2) in the case of degree as centrality measure (Wasserman and Faust 1994; Csardi and Nepusz
4. Exploring News Content Complexity Using Network Analysis

2006). The more centralized a network and thus the higher the centrality value, the less complex the network's structure. Centralization takes values between 0 and 1, and a network is most centralized – and thus least complex – at 1.

**Heterogeneity** We compute the Gini coefficient, used mainly in economics, for analyzing network heterogeneity; measuring to what extent the connections in a network are equally distributed over a number of predefined categories of actors or issues. The normalized formula for the coefficient is

\[
G_N = \frac{n \sum_{i=1}^{n} (2i - n - 1)x_i}{(n-1) \sum_{i=1}^{n} x_i}
\]

Here, \( n \) is the number of links in the network, and \( x_i \) is the number of links towards, from, or towards and from group \( i \) in the ranking (Leydesdorff and Rafols 2011).

For this measure we take into account exogenous factors (node attributes) to classify the actors and issues, and count the frequencies of the links between different categories. To create these categories, we aggregate actors into political and societal groups (e.g. government parties, opposition parties, public institutions), and issues into issue poles based on ideological issue dimensions (see e.g. Van Hoof et al. 2014). A relation between two actor or issue groups is measured by its frequency in the considered media coverage; it is not treated as a binary variable like for the previous measures.

For the actor network, we also take into account the directions of links (i.e. who is a speaker and who is an addressee). The outdegree of an actor group is the sum of its outgoing links. While a high indegree (the sum of all incoming links) would characterize a group that is being addressed or spoken about relatively often, a high outdegree characterizes a group that is often given the chance to speak about or to address actors from other groups. Adam (2008) argues that in mediated networks, speakers use political claims to ‘frame’ other actors, for example by criticizing them or endorsing them. Using the heterogeneity measure for actors, we can analyze to what extent this power to frame others is equally distributed over different actor
groups. If it would always be the same group, it would mean that one group of people or ideas are dominant in a debate, which would make it less heterogeneous and thus less complex.

Since issues are connected to other issues through co-occurrence in the same news story, the separation between indegree and outdegree for issue poles is not possible. For the issue networks, the Gini coefficient is calculated using the degree, meaning the number of different issue poles to which an issue of a particular issue pole is connected in each news story.

The Gini coefficient is at zero when the links in a network are equally distributed between categories. In a complex debate, actor groups or issue poles are connected to all other groups or poles to a more or less equal extent. The more groups or issue poles are connected to the same groups/poles, the less complex the debate and the higher the Gini coefficient, with a maximum of 1. For example, a debate wherein one group is connected a lot to other groups while others are not connected at all results in a high Gini coefficient. The more a media outlet gives space to the voices of various socio-political groups, and the more issues are connected to issues from a diverse range of other issue dimensions, the more the outlet contributes towards informing citizens about a complex political reality (see Napoli 1999).

Because we measure heterogeneity coming from this perspective, we do not only consider how heterogeneous coverage in different outlets is, but also which actor groups and issue poles are given high visibility in each respective outlet. Which actor groups speak to or act most towards other groups in the media? And issues from which poles are often used to contextualize other issues? We call this relative media attention, shown as the percentage of the total number of outgoing links from one group to other groups (for actor groups), or of the total number of links between different poles (for issue poles).

It should be noted that all the complexity indicators we have adapted from graph theory are based on direct links. Though graph theory has also developed a number of complexity indicators looking beyond direct links, such as betweenness centrality, these would be hard to
interpret for mediated networks of political actors and issues in news stories.

All network measures have been calculated using the statistical software R as well as its network analysis package *igraph* (Csardi and Nepusz 2006).

**Results**

**Interconnectedness**

Our first complexity measure is interconnectedness, measured using density (D) – the actual number of connections between actors or issues in a particular outlet, compared to their connections in a fully connected graph. Table 4.3 shows the density for the actor networks.

**Table 4.3 Actor Interconnectedness**

<table>
<thead>
<tr>
<th>Outlet</th>
<th>Nodes</th>
<th>Density</th>
<th>Density (LCD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elite</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Der Standard</td>
<td>206</td>
<td>0.01</td>
<td>0.04</td>
</tr>
<tr>
<td>Die Presse</td>
<td>198</td>
<td>0.01</td>
<td>0.05</td>
</tr>
<tr>
<td>Popular</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kronen Zeitung</td>
<td>262</td>
<td>0.01</td>
<td>0.07</td>
</tr>
<tr>
<td>Heute</td>
<td>145</td>
<td>0.01</td>
<td>0.03</td>
</tr>
</tbody>
</table>

A first look at the interconnectedness of the six actor networks indicates that all networks are connected to an equally sparse extent (D = 0.01). We identified the lowest common denominator (LCD) and found 64 actors that were present in all networks. Based on these actors, we see that the networks across all outlets are still quite sparsely connected. Whereas popular newspaper *Kronen Zeitung* (D = 0.07) is more interconnected than elite newspapers *Die Presse* (D = 0.05) and *Der Standard* (D = 0.04), density in the other popular newspaper *Heute* (D = 0.03) is lowest.

Overall, density is low, between 0.03 and 0.07, meaning that most actors are not interacting with a broad range of different actors, although they might be interacting strongly with individual ones. This points towards media coverage, both in elite and in popular
newspapers, that represents the political process in a relatively simple way, based on the interconnectedness dimension of network complexity.

**Table 4.4** Issue Interconnectedness in density (D)

<table>
<thead>
<tr>
<th>Outlet</th>
<th>Nodes</th>
<th>D</th>
<th>D (LCD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elite</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Der Standard</td>
<td>110</td>
<td>0.07</td>
<td>0.17</td>
</tr>
<tr>
<td>Die Presse</td>
<td>117</td>
<td>0.07</td>
<td>0.17</td>
</tr>
<tr>
<td>Kronen Zeitung</td>
<td>117</td>
<td>0.08</td>
<td>0.20</td>
</tr>
<tr>
<td>Heute</td>
<td>79</td>
<td>0.05</td>
<td>0.07</td>
</tr>
</tbody>
</table>

In Table 4.4, we see that these differences are smaller in the non-LCD network: there, *Heute* has an interconnectedness of $D = 0.05$ for the non-LCD network, compared to the other media that have a higher interconnectedness of $D \geq 0.07$. However, for the LCD network, interconnectedness $D = 0.07$ for *Heute*, compared to $D \geq 0.17$ in other outlets, a greater difference. Furthermore, the other popular newspaper *Kronen Zeitung* has a slightly higher inter-connectedness than the elite papers in its issue networks.

While the differences in interconnectedness between different actor networks are too small relative to the size of the network to show in a graph, this is different for the issue networks, shown in Table 4.4 below. The LCD networks for issues consist of only 54 nodes, and the differences in interconnectedness are visible from their network graphs. We show these in Figure 4.4.
4. Exploring News Content Complexity Using Network Analysis

Figure 4.4 Interconnectedness in the lowest common denominator theme networks for four outlets\(^9\)

*Note: Larger nodes represent nodes with a higher degree, and thicker lines represent more links between nodes.*

Figure 4.4 shows that the *Heute* network, with the lowest density, has visibly fewer connections between nodes than the more densely

\(^9\) Graphs were created using the force based layout algorithm Force Atlas 2 (Jacomy et al. 2014). Parameters were Scaling: 100, Gravity: 150.
interconnected networks from issue coverage in *Der Standard*, *Die Presse* and especially the *Kronen Zeitung*.

In sum, we found that interconnectedness is about the same in one popular and two elite newspapers. The other popular newspaper, *Heute*, stood out in being clearly less interconnected than the other outlets, especially in its issue coverage.

*Centralization*

The second dimension of complexity, centralization (C), gives insight into the extent to which a network is focused on particular actors or issues, thus making it less complex. For this measure, the frequency of connections between actors or issues is not taken into account—it only looks at the variety of these connections.

**Table 4.5** Network Centralization (C)

<table>
<thead>
<tr>
<th>Outlet</th>
<th>Actor Network</th>
<th>Issue Network</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Nodes</td>
<td>C</td>
</tr>
<tr>
<td>Elite</td>
<td><em>Der Standard</em></td>
<td>206</td>
</tr>
<tr>
<td></td>
<td><em>Die Presse</em></td>
<td>198</td>
</tr>
<tr>
<td>Popular</td>
<td><em>Kronen Zeitung</em></td>
<td>262</td>
</tr>
<tr>
<td></td>
<td><em>Heute</em></td>
<td>145</td>
</tr>
</tbody>
</table>

Looking at the actor networks, we can see that elite papers are less centralized (that is, focused on particular actors that address or are addressed by other actors) and thus more complex than popular newspapers. Elite paper *Der Standard* is the least centralized (C=0.12) – the debate is the most varied out of all newspapers – whereas popular newspaper *Kronen Zeitung* is the most centralized (C=0.19).

For the issue networks, we see a different pattern, where differences between papers from the same media type are just as big as the differences between types. Popular newspaper *Heute* (C = 0.19) and elite paper *Der Standard* (C = 0.20) have relatively low centralization, meaning that they lay connections between a broad range of issues in their political coverage. Popular newspaper *Kronen*
Zeitung (C = 0.28) and elite paper Die Presse (C = 0.35) tend to focus on a more narrow range of issues.

When we look at the specific issues that different newspapers focus on (not shown here), we see that both popular and elite newspaper issue networks focus on the issue of corruption – that is, this issue is connected with many other issues. Other central issues in popular media were crime and general economics. The elite media have the issues taxes and government reforms on central positions in their issue networks.

For this indicator, it is actor networks are less centralized and thus more complex, and issue networks are equally complex, but focused on different issues, in elite papers compared to popular papers.

**Heterogeneity and Relative Media Attention**

Heterogeneity, measured using the Gini coefficient, represents the degree of evenness in the distribution of connections between actor or issue groups in a network. For actors, we consider only the outdegree of each group. At a Gini of 0, every actor group is represented in the media as acting towards or speaking to other groups an equal number of times. Figure 4.5 shows both the Gini coefficient as well as the percentage of outgoing claims for each group (relative media attention).

When looking at the Gini coefficients for the different newspapers, no clear difference between media types emerges. Both elite and popular newspapers have one outlet with greater heterogeneity with a lower Gini coefficient (Der Standard, G = 0.69, and Heute, G = 0.68, respectively) and one with lower heterogeneity. Looking at the distribution of claims between groups, we see that popular newspapers give more space to citizens (Kronen Zeitung), the government and the opposition (Heute), whereas elite newspapers tend to give the media, foreign actors, scientific experts and civil society more space.
The Quality of Political News in a Changing Media Environment

**Figure 4.5** Actor Groups Claim Distribution (Outdegree)

*Note: Lower Gini means a more equal distribution of outgoing claims towards other groups and thus higher complexity. The percentages refer to how often each group acts towards or speaks about another group, within each outlet. The bar charts give a visual impression of the evenness of this distribution. Percentages may add up to 101% due to rounding.*

For the heterogeneity of issues, in- and outdegree are the same. Figure 4.6 shows the heterogeneity and the distribution of connections between issue poles (relative media attention), for the issue networks.
Figure 4.6 Issue Poles Claim Distribution (Degree)

Note: Lower Gini means a more equal distribution of claims between different groups and thus higher complexity. The percentages refer to how often each group acts is connected to another group, within claims of one news story, for each outlet. The bar charts give a visual impression of the evenness of this distribution. Percentages may add up to 101% due to rounding.

When looking at issue poles, we do see clear differences in heterogeneity for the public sphere as represented in popular versus in elite newspapers. Popular newspaper *Heute* again has the greatest heterogeneity, and thus the lowest Gini coefficient (G = 0.51) out of all outlets, whereas elite paper *Die Presse* is the least heterogeneous (G = 0.64). As with actor claims, elite and popular newspapers also differ in their issue focus: popular newspapers contextualize issues more
often with green or grey (infrastructure and energy) issues than elite newspapers do. Elite newspapers have a larger focus on liberal issues, such as abortion. Differences within media types are also notable. For example, Der Standard has a comparatively large focus on cosmopolitanism (pro-immigration) issues, and Heute on consensus (e.g. education) and on negative valence issues (e.g. crime).

Do elite and popular newspapers differ regarding their heterogeneity? For actors, we found that heterogeneity differs between individual outlets, but not between media types. For issues, we found that popular newspapers are more heterogeneous and thus more complex than elite newspapers. Also, our results show that elite and popular newspapers differ in terms of which groups they emphasize, that is, in the actor groups they allow to speak, as well as in the ideological issue poles they use to contextualize other issues.

**Summary and Discussion**

In this paper, we argued how complexity measures derived from network analysis and graph theory can be used for measuring news content complexity, and then proceeded to demonstrate their usefulness in a case study on election campaign coverage in popular and elite newspapers.

By adapting measures of traditional network analysis to the analysis of mediated networks, we are providing valuable tools for the evaluation of the political news coverage of different media outlets. These tools represent an important addition to the existing complexity measures, because the latter focus on textual complexity, that is, the structural or grammatical characteristics of texts, such as word difficulty, sentence length, or sentence structure. By contrast, our newly proposed measures capture the complexity of news content, i.e., the complexity of the political debate as it is represented in the news story: how many different political actors acknowledge each other in the debate? Which issues are connected? In how far does the debate focus on specific actors or issues? And how diverse is the debate?
Regarding the three complexity measures we have proposed, we can draw two conclusions. First, as the empirical case study illustrates, our indicators measure different dimensions of content complexity, and the relationship between these dimensions is indeed not linear. For example, the free popular newspaper *Heute* stood out for its low level of interconnectedness, especially for the LCD network of issue coverage. This means that its coverage of actors and issues is less in-depth (fewer connections overall, especially between the same subset of issues present in all media) than that of other newspapers. More concretely, it often discusses single issues or actors within news stories. But it also succeeded in addressing the most heterogeneous ensemble of actors and issues in its overall election coverage. In other words, the news coverage of an outlet can be *simple* concerning the number of connections drawn between issues and actors, but *complex* concerning the diversity of actors and issues discussed. It would be very useful to apply our indicators to a larger sample of news media in order to further probe these relationships between the three indicators.

Second, combining these measurements of the amount of complexity with a descriptive analysis of actor and issue networks (e.g. which actors/issues are at the center?) turned out to be a valuable approach. For example, popular newspapers often connected the issue ‘crime’ from the negative valence issue dimension to issues from other issue dimensions in one news story, while the elite press did so for ‘taxes’ and ‘administration’. ‘Crime’ can be regarded as a soft news issue (e.g. Esser 1999), that is, an easier kind of topic than the hard news that elite media focus on. Network measures alone do not capture this dimension of content complexity or difficulty. They do, however, represent an important addition to the purely descriptive analysis of actors and issues mentioned in the news, in two ways: by capturing the relationships between issues and the relationships between actors and by providing concise measures which greatly facilitate a systematic comparison of the news coverage in different media outlets.
Whereas previous measures remained close to the understandability of news stories based on textual complexity indicators, our measures focus on the complexity of the political debate as represented in the media – which could, in a further step, perhaps be compared to the political debate in reality. However, this points to a remaining research gap: so far, we do not know enough about the relationship between content complexity and understandability: does high content complexity in news also make news harder to understand, or could it actually help people understand issues better (e.g. by giving longer explanations, which were coded in our analysis as extra claims and thus perhaps made the network more complex)? Given that news is a form of communication between media and audiences, communication that needs to be understood (Dahlgren 2000), this relation between complexity and difficulty needs to be explored more, especially from an audience perspective. Our measures of complexity, especially centralization and heterogeneity, bring the concept closer to democratic quality criteria such as diversity than to understandability: more diverse or varied connections lead to higher complexity.

Another important research step would be to explore whether a higher content complexity on the level of mediated networks in media outlets actually does improve the knowledge structure density, and hence the connotative political knowledge, of the respective readers. Previous studies (Eveland et al. 2004; Eveland, Marton and Seo 2004) only gauged the impact of the structural complexity of the presented news texts (e.g. hyperlinks). Here, our network complexity indicators, in particular interconnectedness and centralization, could prove valuable for experiments regarding the impact of news texts of varying content complexity on connotative political knowledge.

On an empirical level, our case study of election coverage showed that the content complexity of elite media does not differ in a systematic way (that is, similarly across all three dimensions) from that of popular media. As an explanatory factor for the complexity of the political public sphere as represented in the media, popular versus elite is not sufficient. Popular newspapers focus on other types of
actors compared to elite newspapers, and use different issues to contextualize their substantial coverage, but this does not necessarily make the content of their coverage less complex. Furthermore, we found substantial differences in content complexity between outlets of the same category, in particular between the two popular papers. The comparatively old and established popular newspaper *Kronen Zeitung* was more similar to the two elite papers than to the free popular paper *Heute* founded in 2004.

This does not imply that our indicators fail to grasp the existing differences in news quality between newspapers. Our findings are validated by the results of other recent studies employing more traditional indicators of news quality such as diversity of mentioned actors (see chapter 3 in this dissertation) or verbal style, focus (privatization) and topics (Magin and Stark 2014) which also conclude that the *Kronen Zeitung* is, in many respects, a ‘quality’ popular paper. Our results thus (re-)confirm the need to be careful when applying labels such as ‘popular’ (or ‘tabloid’) or ‘elite’ without a close look at the concrete content currently provided by news outlets. ‘Elite’ and ‘popular’ are historical categories in a field which is changing rapidly due to the increasing competition online. And any empirical study of news quality should critically reflect its own normative assumptions regarding different aspects of news quality, including content complexity, before choosing and then applying the suitable indicators.

Though our definition of content complexity and the respective indicators were developed to address the question of the quality of political news coverage from a public sphere perspective, it could also be applied fruitfully in contexts outside the political news realm. For news content, our indicators are especially suitable for news about issues that received media attention for an extended period of time, and involved debate by different stakeholders using different sub-issues or angles. The quality of the provided media, including its content complexity, is an important issue in other fields as well, for example for identifying high quality discussions in online fora for citizens’ debates or on political or health questions. In particular if the
link between the content complexity of texts and recipients’ knowledge structure density is confirmed in additional studies, our newly adapted measures would even allow researchers to draw conclusions from the content complexity of texts on any subject to the respective connotative knowledge of their readers.

References


4. Exploring News Content Complexity Using Network Analysis


5. Quantitative Analysis of Large Amounts of Journalistic Texts Using Topic Modelling\footnote{This chapter is derived from an article published in Digital Journalism, 13 September 2015, copyright Taylor & Francis, available online: http://www.tandfonline.com/doi/10.1080/21670811.2015.1093271.}

Carina Jacobi, University of Vienna
Wouter van Atteveldt, VU University Amsterdam
Kasper Welbers, VU University Amsterdam

Abstract

The huge collections of news content that have become available through digital technologies enable and warrant scientific inquiry, thereby challenging journalism scholars to analyze unprecedented amounts of texts. This article proposes Latent Dirichlet Allocation (LDA) topic modeling as a tool to face this challenge. LDA is a cutting edge technique for content analysis, designed to automatically organize large archives of documents based on latent topics, measured as patterns of word (co-)occurrence. We explain how this technique works, how different choices by the researcher affect the results, and how the results can be meaningfully interpreted. To demonstrate its usefulness for journalism research, we conducted a case study of The New York Times coverage of nuclear technology from 1945 to the present, partially replicating the study by Gamson and Modigliani (1989). This shows that LDA is a useful tool for analyzing trends and patterns in news content in large digital news archives relatively quickly.
Introduction

The shift of news media towards online publication and archiving provides journalism scholars with new opportunities for studying journalism. At the same time, understanding the complicated dynamics of this contemporary media landscape requires an ever-larger scale of analysis, with more outlets and more content per outlet. In this article, we show how topic modeling, a relatively new method developed in the computational linguistics field, can help analyse large amounts of text without requiring manual coding, thus reducing the time and costs of such projects. For a general overview of the available methods and pitfalls for topic models, see Grimmer and Stewart (2013). In this article, we focus on one type of topic model, Latent Dirichlet Allocation (LDA) (Blei, Ng and Jordan 2003), and demonstrate its use for journalism research. Even though topic modeling is a promising method for text analysis, with the seminal paper in computational linguistics (the aforementioned Blei, Ng, and Jordan 2003) published around a decade ago, it is just starting to be used in the social sciences. Political scientists, who, just like journalism researchers have both the challenge and the opportunity of newly available online archives of textual data (such as political speeches, legislative documents and social media) have started to use topic models to automatically classify these documents. Notably, Quinn et al. (2010) classify speeches in the US Senate into topics using topic modeling. Lucas et al. (2015) apply topic modeling to different types of documents such as fatwas and social media posts in order to facilitate comparison between countries. Following such studies, we explore if topic modeling can be used to classify journalistic documents into categories that are meaningful to journalism researchers.

LDA, like other topic modeling algorithms, is an unsupervised technique that automatically creates topics based on patterns of (co-occurrence of words in the documents that are analysed. Journalistic texts are thus ‘coded automatically’ for topics, although it is up to the researcher to interpret the results of the model and to set up the analysis in such a way that the results are useful for the study at hand. Thus, the usefulness of the technique for studying
journalism crucially depends on the correspondence between topics and the constructs of theoretical interest. The goal of this article is to introduce LDA to journalism scholars and to provide a practical guide in applying the technique to their own research. Concretely, we will deal with three broad topics:

**What is topic modelling?** The first part of this article will give a brief and mostly non-technical description of LDA.

**How to set up an LDA topic model?** Secondly, we will describe the different parameters of the LDA topic model, and discuss issues of validity.

**Theoretical interpretation.** The last and most important part of the article discusses how LDA topic models relate to theoretical constructs of interest to journalism researchers, especially issues and frames. Using the example of the news discourse on nuclear technology from 1945 to now, we show how LDA topics mostly correspond to the important issues in this discourse, comparing our results to the earlier study by Gamson and Modigliani (1989).

**What is Topic Modelling?**

Topic models are computer algorithms that identify latent patterns of word occurrence using the distribution of words in a collection of documents. The output is a set of *topics* consisting of clusters of words that co-occur in these documents according to certain patterns. In an LDA model, each document may contain multiple topics. Each of the topics has an internal consistency – the words in that topic often occur together in the documents, and/or do not appear much outside of that topic. The researcher determines what this consistency refers to, and thus how the topic can be interpreted (Chang et al. 2009). It is this interpretability that determines whether topic models in general, and LDA specifically, are of use to social scientists.

In the case of journalism research, the collection of topics inferred by the model would ideally resemble a categorization of issues or frames based on substantive theory, for example the issue list used by...
the Comparative Agendas Project that uses categories such as Macroeconomics, Foreign Affairs and Crime to categorize legislative and journalistic texts (Jones and Baumgartner 2005). However, topics are created by the LDA algorithm based on patterns of word co-occurrence in documents, which do not necessarily match theoretical concepts. It seems plausible to equate a topic with the theoretical concept ‘issue’ or ‘theme’, but topics could potentially also represent writing or speaking styles (e.g. many words referring to emotions), events (e.g. a natural disaster) or frames (e.g. immigrants framed as criminals), which, at least theoretically, are also formed through co-occurrence patterns of specific words. For example, in Quinn et al. (2010)’s study of topics in Congressional Record speeches, words such as violence, drug trafficking, police and prison were interpreted as being the topic Law and Crime I: Violence/Drugs. Another topic, Abortion, contained words such as baby, abortion, procedure, and life (Quinn et al. 2010, 217). This study thus interprets topics as issues, whereas DiMaggio, Nag, and Blei (2013, 590-591) refer to ‘voices’ or ‘frames’ when interpreting different topics. However, what exactly topics represent, and if they represent different concepts given different input parameters in the model, is ultimately an empirical question.

To illustrate what topics look like and the interpretation challenges they present, consider the example news story in Figure 5.1. This story, which appeared after the Chernobyl disaster, criticizes the Soviet Union for suppressing news about the event. In the text, each word is highlighted to indicate which topic it was drawn from. These topics were found automatically by the algorithm, but we interpreted and named the topics in a way similar to how one would interpret a factor analysis outcome.\footnote{Factor analysis is a dimensionality reduction technique: given a set of observed variables, a smaller set of factors is calculated that preserve as much information as possible in a lower-dimensional space. This is often used in the field of psychology as a measurement of latent, unobserved causes for certain observations. For instance, if a single factor largely explains the results for a set of questions relating to anxiety, the factor can}
5. Quantitative Analysis of Journalistic Texts Using Topic Modelling

Figure 5.1 Example news story with words from different topics highlighted in the text

As would be expected given the content of the story, many words are drawn from a topic that we interpret as the Nuclear Accidents topic, indicated in dark grey with white letters. This topic includes words such as ‘Chernobyl,’ ‘disaster,’ and ‘radiation,’ but interestingly also contains the words ‘last Friday,’ probably due to the episodic nature of accident reporting. The other main topic, Cold War, is indicated in medium grey and contains words such as ‘Soviet Union’ and ‘Gorbachev’ but also ‘confidence’ and ‘pledge’. Finally, a number of terms such as ‘secrecy’ and ‘peaceful use’ of ‘energy’ are drawn from the Nuclear Research topic, and two words are drawn from two other topics.

This example highlights a number of interesting points about LDA. First, this document is split between two main topics, Cold War and Nuclear Accidents. In a coding scheme forced to have a single topic per document, it would be very difficult to choose the ‘dominant’ topic for this news story, so in our opinion this accurately reflects the nature of the story. Second, you can see that not all words are included in the analysis: most of the words in the story are not used by the algorithm (the text without highlighting), either because they be interpreted as a measurement of anxiety. Similarly, a topic in topic modeling can be interpreted and named based on what the main words have in common.
are non-substantive words such as determiners or prepositions (‘the’ or ‘it’), which we excluded, or because they are too rare (‘Elbe’, ‘perish’), or too common (‘have’ but in this context also ‘nuclear,’ since that was used to select the news stories). Finally, it should be remembered that no a priori coding scheme was used by the computer, so the topics in this document were found completely automatically. The fact that the resulting topics (such as Cold War and Nuclear Accidents) make sense substantively and match the way we would define the subject or topic of news stories – such as the example here – shows that to some extent our notion of ‘topic’ matches the latent classes or co-occurrence patterns identified by the LDA algorithm.

Topics such as ‘Cold War’ and ‘Proliferation’ can be interpreted as issues, but also as ways of framing nuclear weapons. In such cases, a particular perspective on an issue can be signified by specific keywords, which can be seen as framing devices (Matthes and Kohring 2008). In general, if framing devices correspond to specific (latent) patterns of vocabulary use, LDA can capture these classes in specific topics, and as such LDA results can also include the frames used in a corpus of texts.

**How to Set Up an LDA Topic Model**

Before showing a case study where we interpret LDA topics, we will explain how to perform an analysis using this technique. As previously mentioned, LDA processes a collection of documents and clusters the words in these documents into topics in an unsupervised way. As such, it is important to start by considering which documents should be included, and to make sure these documents consist only of the content of news stories – other textual information such as author name or reader comments should be removed.

We will discuss the steps we took when doing the analysis for our case study below. Our data consisted of 51,528 news stories (headline and lead) from the *New York Times* that mentioned nuclear power,
5. Quantitative Analysis of Journalistic Texts Using Topic Modelling

published between 1945 and 2013. We retrieved all news stories from the *New York Times* online archive\(^\text{12}\) that contained the search terms ‘nuclear’, ‘atom’ or ‘atomic’ in the headline or lead.

**Preprocessing: Tokenization, and Lemmatization**

A topic model does not analyse documents directly, but uses a so-called document-term matrix based on these documents. This matrix lists the frequency for each term (word) in each document. The first step in creating this matrix is tokenization, which means splitting the text into a list of words. Although this can be done by splitting on white space and punctuation, there are good word boundary detection tools that recognize acronyms, contractions, et cetera.

It is possible to fit a topic model directly on the document-term matrix containing all the words in the corpus, but it is often better to first reduce the size of this matrix by preprocessing and feature selection. This reduces computing time and improves the results. An important step in preprocessing is stemming or lemmatizing. Stemming is a simple technique where the ending of words is ‘chopped off’, leaving the stem. For example, using the frequently-used Porter stemming algorithm, ‘weaknesses’ becomes ‘weak’, and both ‘failures’ and ‘failure’ become ‘failur’. This technique does not handle irregular conjugations, for example ‘are’ and ‘were’ have different stems. A more powerful (and computationally demanding) technique is lemmatizing. Lemmatizing reduces all words to their ‘lemma’ using a lexicon in combination with regular conjugation rules. Thus, lemmatization reduces both ‘is’ and ‘were’ to their lemma (to) ‘be’. For English, stemming is often sufficient, but for more richly inflected languages such as German or Dutch, lemmatization tends to give better results (Haselmayer and Jenny 2014).

**Feature Selection**

The next step is feature selection. A moderately large corpus can

\(^{12}\) http://developer.nytimes.com
typically contain more than 100,000 unique words, with most of these words occurring in only a few documents. Applying an LDA model on all words in a corpus is both computationally expensive and not very useful, as most words have distribution patterns that do not contribute to meaningful topics. For example, some words are too frequent to be informative – words like ‘the’ and ‘have’ generally occur in every document regardless of topic. A useful technique to filter out words that are too rare or too common words is to use the tf-idf (term frequency—inverse document frequency) score, meaning that words that are either very rare or very frequent get a low score, and can thus be filtered out. Another option is to have a minimal frequency cut-off to filter out the rare words and use a list of common stop words (and/or cap the inverse document frequency) to filter out overly common words.

Lemmatization software is often combined with POS-tagging. POS (part of speech) tags indicate the type of word, e.g. verb, noun, or preposition. For topic modeling, it is often best to only use specific parts of speech, especially nouns, proper nouns and, depending on the task and corpus, adjectives and verbs. This automatically filters out the most common stop words, which tend to be determiners or prepositions (with the exception of common verbs like ‘to be’ and ‘to have’).

For the analysis presented here, we used the lemmatizer and POS-tagger from Stanford’s corenlp suite (De Marneffe, MacCartney, and Manning 2006), and selected all nouns, verbs, adjectives and proper nouns. We filtered out all terms with a frequency of less than 20 and which occurred in more than 25% of documents, and we removed all terms that contained a number or non-alphanumeric characters, yielding a total vocabulary of 8,493 terms.

Choosing Parameters

After creating a sufficiently small and relevant document-term matrix on which to run the model, there are some more choices the researcher needs to make before running the model. Although one of
the main advantages of a topic model is that no a priori coding schemes need to be supplied, there are certain parameters that need to be set. In particular, the number of topics (K) needs to be specified, which indicates into how many topics the LDA model should classify the words in the documents. There is no default or simple rule of thumb for this parameter. The trade-off is comparable to factor analysis: the goal is to describe the data with fewer dimensions (topics) than are actually present, but with enough dimensions so that as little relevant information as possible is lost. We first discuss and use the perplexity measure, which is a commonly used computational indication for the correct amount of topics (Blei, Ng, and Jordan 2003). However, we stress that this should be only used to make an initial selection of models with acceptable amount of information loss, and that interpretability of topics is a more important criterion for social science purposes.

A second parameter is the \( \alpha \) hyper-parameter, which affects how many topics a document can contain\(^{13} \). A common default value for the alpha is 50 divided by the number of topics. Substantively, a lower alpha leads to a higher concentration of topic distributions within documents, meaning that documents score high on a few topics rather than low on many. Accordingly, if the goal is to assign one or a few topics per document then it makes sense to use a low alpha (Kim et al. 2014).

**Perplexity** From a computational perspective, a good indication of the right number of topics is that number with which the model best predicts the data. This is comparable to goodness-of-fit measures for

\(^{13}\) Technically, the alpha hyper-parameter controls the concentration of the dirichlet distribution regarding the distribution of topics over documents. In Bayesian statistics, a hyper-parameter is a parameter that controls distributions such as the dirichlet distribution. The term hyper-parameter is used to distinguish them from the parameters of the topic model that is the result of the analysis. For a good explanation of the role of hyper-parameters, we suggest the introduction to the dirichlet distribution by Frigyik, Kapila, and Gupta (2010).
statistical models. For topic models such as LDA, a commonly used indicator is perplexity, where a lower perplexity indicates a better prediction (Blei, Ng, and Jordan 2003). To calculate the perplexity, we first train an LDA model on a portion of the data. Then, the model is evaluated using the held-out data. This routine is repeated for models with different numbers of topics, so that it becomes clear which amount leads to the lowest perplexity.

Figure 5.2 shows the perplexity of different models for our data. We trained the LDA models using 30,000 of the 48,604 documents, and then calculated the perplexity of each model over the remaining 18,604 documents. We varied the number of topics used: 2, 5, 10, 25, and then on to 250 in steps of 25. The results show that perplexity decreases as the number of topics increases, implying that a model with more topics is better at predicting the held-out data. At around 25 and 50 topics, the decrease in perplexity for additional topics becomes notably less. This is one way to interpret the right number of topics, similar to the interpretation of the elbow in the scree plot of a factor analysis. The other way is to look at the number of topics with the lowest perplexity. We can see that this point will be somewhere beyond 250.

However, having the ‘right’ number of topics in a mathematical sense does not say anything about the interpretability of the topics produced. In fact, mathematical goodness-of-fit measures and human interpretation may lead to contradicting conclusions on the best number of topics, given that especially with a high number of topics, the computer algorithm may find nuances that are not semantically meaningful to humans (Chang et al. 2009). Also, as we use topics to answer substantive questions about the documents we study, it is important that the topics that result from the analysis contribute

---

14 A goodness-of-fit measure describes how similar the predicted or expected values of a model are to the actual observed values. An example is the $R^2$ measure in linear regression, which indicates what proportion of variance of the dependent variable is explained by the independent variables.
towards answering these questions, instead of providing the best prediction of the data. For the analysis presented in this article, we thus looked both at perplexity and interpretability when deciding on the number of topics to use. Judging from the perplexity, a good choice is probably between 25 and 50 topics. However, to facilitate comparing our results with Gamson and Modigliani (1989), we will first use a simpler model with $K = 10$ topics, and then show the differences between this model and a model with $K = 25$ topics.

![Perplexity of LDA models with different numbers of topics and alpha.](image)

**Figure 5.2** Perplexity of LDA models with different numbers of topics and alpha.

*Note: The line graph shows how perplexity decreases (and model fit thus increases) as the number of topic increases. The number of topics that corresponds to a great change in the direction of the line graph is a good number to use for fitting a first model. For example, $K=25$ for our model with alpha $= 50/K$.*

**Alpha** Regarding the alpha hyper-parameter, since the data used for our analysis covers a long history of textual data concerning an issue that involves various events and viewpoints, it makes sense to define several clearly distinguishable topics. In addition to better scores for the lower alpha, this is a substantive argument for us to use the lower alpha of five divided by $K$ instead of the default.
Tool Support for LDA

The easiest way to get started with LDA is through the open-source statistical package R.\textsuperscript{15} Although specialized software for topic models is available, such as MALLET (McCallum 2002) or the Stanford Topic Modeling Toolbox (Ramage et al. 2009), an advantage of using R is that it is a statistical package that many social scientists already use for other analyses. However, LDA works the same no matter the software used to run the model. The \texttt{tm} package in R can automatically generate the document-term matrix from texts and includes options for stemming and feature selection (Meyer, Hornik, and Feinerer 2008). The \texttt{topicmodels} package can directly fit an LDA model from the document-term matrix object created by \texttt{tm}. For more sophisticated preprocessing, we use the Stanford \texttt{corenlp} suite (De Marneffe et al. 2006), which contains a collection of modules for grammatical analysis. For our case study, we uploaded our documents in the web-based content analysis toolkit AmCAT\textsuperscript{16}. For the preprocessing and the analysis itself, we used the statistical computing environment R, that can connect to AmCAT directly to retrieve the documents there\textsuperscript{17}. AmCAT uses the \texttt{xtas} extensible text annotation suite to automatically perform the preprocessing in a scalable manner (De Rooij, Vishneuski, and De Rijke 2012)\textsuperscript{18}.

Nuclear Topics: 1945 - 2013

In order to demonstrate the use of LDA to explore the topics in a given set of documents, and show the change in these topics over time, we have performed LDA on news stories dealing with nuclear technology from the New York Times. The famous study by Gamson and Modigliani (1989) shows how the framing of the issue ‘nuclear

\textsuperscript{15} http://www.r-project.org
\textsuperscript{16} http://amcat.nl
\textsuperscript{17} See http://github.com/amcat/amcat-r for the relevant R code. The R scripts that were used for our analysis can be downloaded from http://github.com/AUTHOR/corpus-tools.
\textsuperscript{18} See http://github.com/AUTHOR/xtas for the xtas modules for corenlp and other lemmatizers.
power’ in the news changed over time since 1945. We were interested to see whether the topics found by performing an LDA over news stories from this period shows similar changes over time. Additionally, we extended the research period to 2013, to include more recent coverage of nuclear technology.

The question is thus whether the change in culture surrounding nuclear issues found by Gamson and Modigliani is expressed in a change in word use over time that is captured by LDA. Gamson and Modigliani (1989) identified seven ‘packages’ or frames in newspaper and television coverage of nuclear energy between 1945 and 1989: Progress, Energy Independence, Devil’s Bargain, Runaway, Public Accountability, Not Cost Effective, and Soft Paths (Gamson and Modigliani 1989, p 24-25). If we compare the outcome of an LDA analysis with the results of Gamson and Modigliani’s study, to what extent do we find similar results? Note that for a number of reasons we do not expect perfect correspondence between the LDA topics and Gamson and Modigliani’s packages, even if we disregard the difference between manual and automatic analysis. First, these packages were identified by examining not only the text of news stories, but also images and cartoons, with a focus on editorial content, whereas our analysis is performed on the lead paragraphs of news stories only. Second, whereas Gamson and Modigliani only analyse nuclear power, our investigation deals with the whole coverage of nuclear technology, including nuclear weapons. Finally, our investigation covers the post-Cold War period as well, while Gamson and Modigliani of course only analyse the discourse until the 1980s. Nonetheless, it is interesting to compare our findings to theirs since it shows to what extent the results of an automatic topic modelling approach compare to those of a well-known and very thorough manual analysis.

As discussed above, for reasons of comparability to Gamson and Modigliani’s seven packages, we will first focus on an analysis using 10 topics. Table 5.1 shows the topics that resulted from this analysis, including our interpretation and the 10 words that represent a topic.
Table 5.1 LDA results on U.S. Nuclear discourse, 10 topics

<table>
<thead>
<tr>
<th>Topic</th>
<th>Interpretation</th>
<th>Most representative words</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Topics with temporal patterns</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Research</td>
<td>atomic, Energy, WASHINGTON, scientist, energy, bomb, Commission, United, research, weapon</td>
</tr>
<tr>
<td>3</td>
<td>Cold war</td>
<td>United, States, Union, Soviet, soviet, weapon, arm, missile, President, treaty</td>
</tr>
<tr>
<td>7</td>
<td>Proliferation</td>
<td>Iran, United, North, Korea, program, weapon, States, official, country, China</td>
</tr>
<tr>
<td>8</td>
<td>Accidents/Danger</td>
<td>plant, power, reactor, Island, Nuclear, accident, Commission, official, waste, safety</td>
</tr>
<tr>
<td><strong>Continuous topics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Weapons</td>
<td>test, submarine, Japan, first, Navy, year, explosion, missile, ship, bomb</td>
</tr>
<tr>
<td>9</td>
<td>Nuclear Power</td>
<td>power, plant, company, year, energy, percent, utility, cost, Company, reactor</td>
</tr>
<tr>
<td>10</td>
<td>US Politics</td>
<td>war, President, weapon, Mr., year, military, policy, world, Reagan, House</td>
</tr>
<tr>
<td><strong>Irrelevant topics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Summaries</td>
<td>New, new, year, government, official, York, people, business, President, state</td>
</tr>
<tr>
<td>4</td>
<td>Book Reviews</td>
<td>week, life, book, man, woman, John, year, New, family, University</td>
</tr>
<tr>
<td>6</td>
<td>Films &amp; Music</td>
<td>Street, West, Theater, Mr., Sunday, East, show, New, tomorrow, p.m.</td>
</tr>
</tbody>
</table>
To facilitate interpretation we also made a topic browser\(^{19}\), which is a tool to interactively explore the results of a topic model (Gardner et al. 2010). Our topic browser features two extra pieces of information in addition to top words. One is the top documents assigned to a topic in which the topic words are highlighted. The other is a semantic network, or semantic map, which visualizes the co-occurrence of top words, thus showing topic coherence and facilitating interpretation. For our discussion, we categorized the topics into: (a) topics that show a pattern in their use over time; (b) topics that are more or less continuously present; and (c) three topics that turned out to be irrelevant for our case study.

**Topics that show some pattern in their use over time.** Firstly, we found a number of topics that have a strong temporal dimension, that is, they are strongly present in the news in some years or decades, but not in others. In that respect, these topics are most similar to the shifting packages found by Gamson and Modigliani. Figure 5.3 shows the change in occurrence over time for the four topics discussed below, which are discussed in chronological order.

News stories in which topic 1, which we labelled *Research*, deal with research on nuclear technology, including both energy research and nuclear weapons research. This topic is most strongly present in the early part of the data set, and its usage sharply decreases over time, especially from the 1980s onwards. In terms of temporal focus, and in the focus on possibilities created by nuclear research, this topic is comparable to Gamson and Modigliani’s Progress package, although the latter did not include nuclear weapons research as their research was focused on nuclear energy.

In topic 3, *Cold War*, the words ‘United’, ‘States’, ‘Soviet’, ‘Union’ and ‘weapon’ immediately suggest that this is a topic about the US–Soviet conflict. Lower-ranking words include variations on the ‘weapon’ theme as well as more diplomatic terms such as ‘agreement’ and

\(^{19}\)The topic browser can be found at http://rpubs.com/Anonymous/78706
'proposal'. The topic occurs most frequently between the mid-1950s and the mid-1980s, with a peak in the early 1960s that can be easily identified as the Cuban missile crisis.

Figure 5.3 Occurrence of topics that have a strong temporal component

And finally, the last topic in this category is topic 8, Nuclear Power–Accidents/Danger. Although the top words in this cluster, ‘plant,’ ‘power,’ and ‘reactor’, are not very informative, peripheral words like ‘accident,’ ‘safety,’ and ‘radioactive’ show how to interpret it. Indeed, Chernobyl and Three Mile Island are both included in this topic, which shows peaks in news attention in both 1979 and 1986. News stories on smaller nuclear accidents or discussions on nuclear hazard in general also contain this topic. Interestingly, although this topic shows a small peak around the 2011 Fukushima reactor, most of the coverage of that event is classified in topic 5 (nuclear weapons), discussed below. This topic is most closely related to Gamson and Modigliani’s Runaway and Devil’s Bargain packages, as it focuses on the negative qualities of nuclear energy, and both the topic and the Runaway frame occur in roughly the same time period.

Finally, the news stories in topic 7, Nuclear Proliferation, deal with
nuclear weapons in countries such as North Korea, Iran and Pakistan, and US actions or policies against the possession of these weapons. Peaks of attention for these topics occur in the 1990s and especially the 2000s, with events in North Korea and later Iran as triggers. Although the topic occurs after the period covered by Gamson and Modigliani, semantically it resembles their *Runaway* package, in the sense that the technology is no longer under control and now poses a danger to its very inventors.

**More or less continuously present topics** These topics do have some fluctuation in their occurrence over time, but this fluctuation shows no clear trend.

In topic 5, *Nuclear Weapons*, the words associated with the topic seem to be mostly related to development and testing of nuclear bombs, especially nuclear submarines. This topic is most strongly present in the 1950s through the early 1970s, with a peak in 2011 after the Fukushima incident. This latter peak is possibly a confusion caused by the prevalence of ocean-related words in both the tsunami coverage and the discourse on nuclear submarines.

Topic 9, *Nuclear Power* is a second cluster with words related to economics. Similar to topic 8, ‘power’ and ‘plant’ form the top words of a cluster. However, this time the peripheral words show a different focus: ‘company,’ ‘corporation,’ ‘year,’ and ‘percent’ suggest that these news stories are about nuclear energy as a business, and report on the earnings of companies that operate in this sector. Over time, we see a peak in the 1970s (the oil crisis) and in the 1980s.

The final relevant topic, topic 10, *US Politics*, concerns the policies of the United States on its own nuclear weapons and defence. The main peak for this topic is in the early 1980s, with the Rearming America programme.

**Irrelevant topics** These topics have nothing to do with nuclear power in terms of their content, but appear in our results anyway
because we included all news stories mentioning the word ‘nuclear’, including book reviews and news summaries. In topic 2, news summaries and items from the section ‘Inside The Times’ are clustered together as one topic. These tend to focus on New York and on politics (local, domestic or foreign). However, they do not deal with nuclear issues directly, so we discard this topic. Topic 4 consists of short book reviews, of which one or more use words related to the nuclear topic. Again, these do not deal with nuclear issues in the news directly. Lastly, topic 6 represents news stories on films or concerts, some of which have to do with nuclear power, others show up in stories that mention the nuclear issue elsewhere.

As this overview shows, not all topics contain useful results, and the topics are not ordered in a way that makes it easy to distinguish the useful from the non-useful (as usefulness is of course something determined by the researcher, not the computer). However, the irrelevant topics were clearly distinguishable, which makes it easy to discard them altogether or ignore them in further analysis. Although it might seem annoying at first that such topics are also generated by the analysis, this is actually quite useful. Since most data sets contain a degree of noise (such as book reviews or sports results), LDA can be used as an inexpensive way to locate those news stories that are relevant for answering a particular research question from a larger sample.

Our purpose here however was to compare the topics we found using LDA with the outcome of the framing research by Gamson and Modigliani (1989). Compared to the frames or interpretive packages found in their analysis, the topics of our LDA analysis seem to be more concrete and specific. Similar to Gamson and Modigliani’s study, we found that topics change over time, but not all of them increase or decrease in a linear way. Also, topics seem to either cluster a number of related events together (nuclear proliferation talks, nuclear accidents) or represent issues that are continuously present over a longer period of time (the economics of nuclear energy, nuclear weapons tests). It is not possible, however, to deduct a particular viewpoint or frame from a topic directly— for example, we found no clear ‘anti-nuclear’ cluster,
whereas Gamson and Modigliani found multiple frames that are critical of nuclear energy. It is quite likely that the coverage of nuclear accidents and danger is predominantly covered from an anti-nuclear perspective, but even in this case there is a clear difference between the ‘issue’ or event being covered (nuclear accidents) and the frame with which it is covered. That said, for some of the topics, such as the Research and Accidents topics, we do see that the temporal pattern is similar to that identified for the Progress and Runaway packages, respectively.

**Granularity: How Many Topics?**

To see whether increasing the number of clusters helps find word patterns that are more fine-grained and more frame-like than those representing issues, we will explore what happens if we increase the number of topics from 10 to 25. Changing the number of topics (K) changes what is called the granularity, or level of detail, of the model. The higher the granularity, the more detailed the analysis. Using a larger number of topics implies higher granularity: each topic then represents more specific content characteristics.

If a model with 25 topics is compared to a model with 10 topics, some of the topics in the model with 25 topics tend to blend together in the 10-topic model. The topics in the model with 50 topics can thus be considered to represent smaller grains of the 10 topics. As an example, consider the events in Chernobyl and Three Mile Island. Both events are clearly separated by time, space, the actors involved and various circumstances. Yet the two are related by the nature of the event, and thus in vocabulary used: a malfunction of a nuclear reactor, with awful consequences to health and the environment. In a model with many topics, these events can be distinguished in different topics, whereas in a model with fewer topics these events are blended together, representing a broader theme. In our data, we saw this specific example in a comparison of models with 10 and 25 topics.
**Table 5.2** Most similar K = 25 topics for each relevant K = 10 topic

<table>
<thead>
<tr>
<th>Topic</th>
<th>Similarity</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1: Research</strong></td>
<td></td>
</tr>
<tr>
<td>9: Research</td>
<td>0.68</td>
</tr>
<tr>
<td>8: Universities</td>
<td>0.39</td>
</tr>
<tr>
<td>17: Scientific development</td>
<td>0.35</td>
</tr>
<tr>
<td>21: Nuclear Weapon Materials</td>
<td>0.27</td>
</tr>
<tr>
<td><strong>3: Cold War</strong></td>
<td></td>
</tr>
<tr>
<td>7: Rearming America</td>
<td>0.61</td>
</tr>
<tr>
<td>18: Cold War</td>
<td>0.61</td>
</tr>
<tr>
<td>20: NATO</td>
<td>0.50</td>
</tr>
<tr>
<td><strong>8: Accidents/Danger</strong></td>
<td></td>
</tr>
<tr>
<td>24: 3 Mile Island</td>
<td>0.70</td>
</tr>
<tr>
<td>15: Power Plant Construction (Shoreham)</td>
<td>0.62</td>
</tr>
<tr>
<td>2: Chernobyl &amp; Fukushima</td>
<td>0.32</td>
</tr>
<tr>
<td><strong>7: Nuclear Proliferation</strong></td>
<td></td>
</tr>
<tr>
<td>6: Iran</td>
<td>0.63</td>
</tr>
<tr>
<td>4: North Korea</td>
<td>0.58</td>
</tr>
<tr>
<td>14: Iraq</td>
<td>0.42</td>
</tr>
<tr>
<td>3: India &amp; Pakistan</td>
<td>0.25</td>
</tr>
<tr>
<td>21: Fissile Materials</td>
<td>0.25</td>
</tr>
<tr>
<td><strong>5: Nuclear Weapons</strong></td>
<td></td>
</tr>
<tr>
<td>11: Nuclear Submarines</td>
<td>0.53</td>
</tr>
<tr>
<td>2: Chernobyl &amp; Fukushima</td>
<td>0.39</td>
</tr>
<tr>
<td>19: Nuclear Weapons</td>
<td>0.31</td>
</tr>
<tr>
<td><strong>9: Nuclear Power</strong></td>
<td></td>
</tr>
<tr>
<td>23: Nuclear Power</td>
<td>0.52</td>
</tr>
<tr>
<td><strong>10: US Politics</strong></td>
<td></td>
</tr>
<tr>
<td>1: Nuclear War Threat</td>
<td>0.68</td>
</tr>
<tr>
<td>5: US Politics</td>
<td>0.46</td>
</tr>
<tr>
<td>19: Nuclear Weapons</td>
<td>0.36</td>
</tr>
<tr>
<td>12: Protests</td>
<td>0.28</td>
</tr>
</tbody>
</table>

*Note: Topics from the 10-topic model (K = 10) are set in italics, with the most similar topics from the 25 topic model (K = 25) listed beneath each K = 10 topic. Some K = 25 topics occur twice when they were similar to more than one K = 10 topic. Similarity is based on whether topics occur in the same documents*
(calculated as cosine similarity). Example: topic 3 from the K = 10 model, interpreted as the Cold War topic, is similar to three topics from the K = 25 model; topic 7 on Reagan and Rearming America (sparking the 1980's arms race); topic 18 on the Cold War; and topic 20 on the NATO.

Table 5.2 gives an overview of the relevant topics from the 25-topic model. Each ‘detailed’ (K = 25) topic is listed below the ‘broader’ (K = 10) topic it resembles most. This similarity is computed by determining in which documents each K = 10 and K = 25 topic occurs, and then calculating the cosine similarity of these occurrence vectors. So, two topics with perfect similarity (1.0) would occur in exactly the same (relative) frequency in all documents. For instance, topic 24, which we labelled 3 Mile Island, is most similar to topic 8 (with cosine similarity 0.70).

In the K = 10 model, coverage of Chernobyl is clustered together with coverage of the Three Mile Island incident, to form a cluster about nuclear accidents (topic 8), and coverage of Fukushima was split between this topic and the nuclear weapons topic (topic 5). However, in the K = 25 model, different connections between these events are found, and Chernobyl and Fukushima end up in the same cluster, while Three Mile Island gets its own cluster. Finally, topic 8 (K = 10) is highly similar to topic 15 (K = 25), which deals with power plant construction and especially the Shoreham nuclear power station. This power station was constructed on Long Island in the 1980s, but never actually used, as local residents objected in the wake of the Chernobyl disaster. As shown in Figure 5.4, these three constituent topics in the K = 25 model together trace the K = 10 Accidents/Danger topic (the per-year correlation between the K = 10 topic and the sum of the K = 25 topics is 0.95, p<0.001), but each are focused around a specific time point.

**Discussion: Best Practices**

This article showcased a relatively new tool, Latent Dirichlet Analysis, or LDA. LDA is an unsupervised topic modelling technique that automatically creates ‘topics’, that is, clusters of words, from a collection of documents. These topics may represent issues that recur
over time, related events or other regularities in news stories.

![Figure 5.4](image)

**Figure 5.4** Occurrence over time of detailed (K = 25) topics that constitute the Accidents/Danger topic from the K = 10 model

We think LDA can be a valuable tool in any large-scale content analysis project. For preliminary analysis, LDA can very quickly give a rough overview of what kind of topics are discussed in which media or time periods.

However, the best way of proving the statistical, internal and external validity of LDA and of topic models in general is still under discussion (Chang et al. 2009; DiMaggio, Nag and Blei 2013; Ramirez et al. 2012). We would advise journalism scholars to start by making a perplexity plot for different numbers of topics, and then look at the models where the perplexity decrease drops off. Furthermore, the researcher should also manually inspect these topic models, and for each topic decide on the correct interpretation by looking at the words in the topic, in which media and time frames it occurs, and also at the documents that are most indicative of a specific topic. For this purpose
we advise using a topic browser\textsuperscript{20}. Manual interpretation reveals which topics are closely related to the theoretical quantities the researcher is interested in. Also, the researcher can decide to combine multiple topics that are semantically related and/or to remove irrelevant topics such as the book reviews identified in our case study. This should be followed by a formal internal validity evaluation of these topics by checking a sample of automatically coded news stories or by comparing to a sample of manually coded stories. See DiMaggio, Nag and Blei (2013) for a more elaborate discussion of validity checks of LDA topic models. Even though the validity may be insufficient for completely automatic analysis, it can be used as a form of semi-automatic coding by having the manual coder check whether the coding is correct, which is much quicker (and cheaper) than fully coding each document.

Furthermore, even if the topics cannot be immediately used to answer the substantive research question, they can enhance the subsequent manual or automatic content analysis in numerous ways. First, by inductively showing which topics occur, it can help the researcher create or improve the codebook by suggesting new codes and examples and by showing which codes might need to be combined or split up. Second, if the researcher is interested in doing automatic keyword analysis, the word lists per topic can offer inspiration for keywords and synonyms that might otherwise be left out. Third, LDA can be used to quickly find good example documents for infrequent topics or documents that use multiple topics, such as the example document in Figure 5.1. This can be used for manual coder training and evaluating the codebook, but also for creating the training data for subsequent machine learning, where rare topics usually give the worst performance because of the lack of training documents. Finally, as shown by the three ‘irrelevant’ topics that were derived from our analysis, LDA can be used to filter out categories of texts that are not relevant from an overall sample.

\textsuperscript{20} Our R script for creating a topic browser is available at http://github.com/vanatteveldt/topicbrowser.
The comparison in this article also shows two limitations of LDA for analyzing journalistic texts. First, not all topics represent substantive word clusters, but also other consistent patterns of co-occurrence such as type and writing style. This is most dramatically shown by ‘irrelevant’ topics such as book reviews, but can be considered beneficial by allowing the researcher to quickly discard such clusters of documents. However, it is also possible that shifts in word use, such as from ‘atomic’ to ‘nuclear,’ caused documents to belong to different topics in the 1950s as compared to the 2000s even though they are substantively similar. The same would hold for writing style differences between different media and especially different formats (e.g. print media versus TV or online). Dynamic Topic Models (Blei and Lafferty 2006) and Structural Topic Models (Roberts et al. 2014) are two extensions of LDA that explicitly deal with shifts over time and between groups of documents, respectively.

Second, the topics identified in this study did not approach what we could call a ‘frame’ in the sense that Gamson and Modigliani (1989) or Entman (1993) would use this concept – as coherent interpretative packages. Similarly, the topics did not represent explicit valence or sentiment – there were no clear pro- or anti-nuclear topics. Although attempts to combine LDA and sentiment analysis in one model have been made (Lin and He 2009), since news coverage is difficult for automatic sentiment analysis in general (Balahur et al. 2013), such methods will likely not yield sufficiently valid results for journalism study in the near future. Following DiMaggio, Nag, and Blei. (2013, p.603), we suggest further research can combine LDA as described here with frame and sentiment analysis using other methods, e.g. using machine learning (e.g. Burscher et al. 2014).

By showcasing LDA and by showing some best practices for running and interpreting model results, this article contributes to the adoption of topic modelling in the practice of journalism research, which is a useful technique for every digital journalism scholar to have in their toolbox to deal with the very large data sets that are becoming available. Although the burden of making sense of the results is still on the
researcher, LDA offers a quick and cheap way to get a good overview of a collection of documents so that substantive questions can be answered immediately, especially about broad patterns in topic use over time or between media. Additionally, it is very helpful for performing preliminary analysis before venturing on a more traditional (and expensive) automatic or manual content analysis project.

References


Lucas, Christopher, Nielsen, Richard A., Roberts, Margaret E., Stewart, Brandon M., Storer, Alex, and Tingley, Dustin (2015). Computer
assisted text analysis for comparative politics. Political Analysis 23, 254-277.


6. Discussion and Conclusion

In modern democracies, mass media are essential for informing people about the world. They offer political news that informs people about the policies, performances and plans of the politicians that represent them, and about the issues in public affairs that should be solved. Political news may furthermore help people to create their own view about political issues by presenting them with various opinions, and vote for the politicians or parties that represent their opinion best (Lau and Redlawsk 1997). (Political) news high in quality is thus important for a well-functioning democracy (Sparks 2000; Strömbäck 2005). But what does ‘high in quality’ mean? And do all outlets have to produce political news with the same characteristics, if their goal is to make news high in quality – regardless of their audience?

News quality standards derived from the electoral model of democracy are often used in studies on news performance (e.g. Esser 1999; Sparks 2000; Donsbach and Büttner 2005). This model offers a strict set of criteria for ‘quality’ news, irrespective of its audience. It demands detached, relevant and diverse political information. The participative model, favoured by supporters of popular news (e.g. Örnebring and Jönsson 2004; Friedrich and Jandura 2012), instead allows an emotionalized and personalized political discourse, as long as peoples’ participation in democracy is secured.

The goal of this dissertation was to explore the continued suitability of (political) news quality criteria derived from the electoral democratic model, both theoretically and empirically (do ‘quality’ elite media actually perform better than popular media?). Additionally, a methodological goal was to explore how new measures and (digital) methods developed outside the field of
political communication can contribute to analyzing news content and quality in the current fragmented media environment.

For these purposes, I used insights from studies on popular news and news reception (which tend to favour the participative model) and discussed the relevance of these findings in the context of the current digitalizing and popularizing news environment and the seemingly growing political disinterest. I investigated how different Austrian media outlets perform when held to news quality standards derived from the electoral democratic model, focusing on political news during the 2013 Austrian National Election Campaign. I developed a new measure for to measure content complexity, one aspect of news quality, and used text mining to measure diversity, leader focus and emotionalization, three further aspects. Lastly, I compared the results of an automatic analysis using topic modeling with that of an existing study, to demonstrate what topic modeling can contribute.

The findings of these studies are summarized below.

Summary

Chapter 2 focused on the first research question, Given recent findings and theoretical insights regarding the democratic functions of popular news, how relevant are quality criteria derived from the electoral model of democracy still? I distinguished news quality in general, and political news quality by electoral democratic and by participatory democratic standards, and explained the thematic and pragmatic dimensions of news quality. Then, I summarized and categorized insights from two strands of research on the role of popular(-ized) news in democracy, one critical and one supportive.

Based on this discussion, I argue that some of the news quality criteria derived from the electoral democratic model are still relevant for evaluating political news quality during an election campaign. This model provides clear criteria for ‘quality' news in this specific context—news that informs its audience about politics. But it should be more lenient towards popular news elements such as an emotionalized and simplified news style, since these may increase news understanding and motivation to participate (including voting).
for certain audiences. These elements, such as an emotionalized and simplified news style, may help increasing political interest, political knowledge and motivation to participate (including voting) among certain audience groups, which is necessary given the current media environment characterized by an abundant, but individualized news supply that makes it more difficult for lesser-interested audiences to encounter information about politics in the media.

To create a set of news quality criteria that incorporates this insight, I combined news quality criteria derived from the electoral and the participative democratic model into a new set of criteria that includes elements which can make political news more interesting and easy to understand for its intended audience as aspects of quality. These criteria can be used to evaluate political news quality during an election campaign. They form the theoretical contribution of this dissertation.

Chapter 3 and 4 provided empirical results on how different media types score on electoral democratic quality criteria, thus answering the second research question, *What are the differences and communalities of political coverage in popular and in elite media in Austria, in terms of their (electoral democratic) quality?*

In chapter 3, I compared the quality of political news between the online and the print edition of Austrian newspapers. I argued that the unlimited space and increased interactivity could contribute to higher news quality in online editions compared to print. However, the faster news cycle and increased competition online might instead make producing news high in quality more difficult. I investigated the amount of news, its (party) diversity, low leader focus and low emotionalization as aspects of quality. Results showed that online editions publish more political news, and in this political news, show higher diversity and lower emotionalization, than print editions. Leader focus was however also stronger in online editions than in their print equivalents. Political news in online editions thus seemed to be higher in quality compared to that in print editions. But this picture changed when considering popular and elite outlets separately: while elite papers performed better online, popular
newspapers provided fewer political news stories and showed a stronger leader focus in their online editions than in their print editions. Moreover, the orientation function of popular newspapers seemed to be at peril in the online editions due to smaller opinion sections. I concluded that the gap in political news quality between elite and popular newspapers is rather small in print but much wider online, to the disadvantage of popular newspapers.

Chapter 4 complemented the previous chapter by studying a quality indicator related to diversity, content complexity, in-depth. Using complexity indicators from network analysis (graph theory), I split complexity up into three dimensions: interconnectedness, centralization, and heterogeneity, which I used to measure the complexity of actor and issue networks constructed from political news coverage. I found that content complexity is not generally higher in elite newspapers than in popular newspapers, but varies between individual outlets, especially between the free and the paid popular newspaper. However, elite and popular papers emphasize different actor groups who speak to or act towards others in the news, and they emphasize different issues.

Given the findings of chapter 3 and 4, the answer to RQ2 is that we cannot simply say that news in Austria during the national election campaign of 2013 showed lower quality in popular papers than in elite ones, at least in print. I will later return to the question whether this is because of the specific national (or temporal) context, or whether I could have found something that can be generalized to other contexts as well. Still, this finding in itself contributes more nuance to the debate about popular news versus ‘quality’ (elite) news, by showing that the two types are not always that different in quality. To put it bluntly, audiences for elite news should not consider ‘their’ news so much better than the ‘other’ news, that of ‘tabloids’ – at least for the print editions.

The final aim of this dissertation is to contribute to the development of new measures and methods for news quality research. Chapter 4 introduced a new measure for content complexity, while chapters 3
and 5 added to the use of new automatic methods for quantitative content analysis in communication studies.

In chapter 4, I showed how network analysis measures derived from graph theory can be used to measure three dimensions of content complexity (interconnectedness, centralization and heterogeneity) in news content, displayed as networks of actors and issues. The interconnectedness dimension is conceptually related to difficulty: more complex news has more explicit connections between its elements, and can thus be perceived as denser in information, which is presumably harder to understand. Centralization looks at how strongly these networks are centred on particular actors and issues. As such, centralization is conceptually close to diversity. This is also the case for heterogeneity, which measures how strongly the news is focused on particular actor groups who act towards or speak about other groups, or on particular ideological issue poles used for contextualizing issues from other poles. Measuring content complexity in this way proved useful not only as an indication of the complexity of actor and issue per se, but also in combination with a descriptive analysis of the distribution of links within the networks. These show which actors get the chance to speak towards or about others, and which issues are often used by outlets to contextualize other issues, thus giving insights into the actor or issue emphases of individual news outlets.

Furthermore, I explored the use of automatic content analysis methods. Search strings were used in chapter 3 to measure relevant indicators for news quality (by standards derived from the electoral democratic model) automatically, showing that these can be measured validly without relying on manual coding. Chapter 5 presented the topic model or word-clustering algorithm Latent Dirichlet Allocation (LDA), and demonstrated how it can be used to get an overview of the thematic content of a large set of news stories. I used a replication of a previous study to show how its results are similar to that of a manual content analysis.

Methodologically, this dissertation thus offered a new way to measure an aspect of news quality, content complexity. Further studies could
use this measure in combination with existing textual complexity measures, and thus provide a more complete overview of complexity in the news over time. Also, I argued that in a digitalized and fragmented media environment, automatic content analysis methods are an important supplement to manual ones for their relatively high speed and thus low costs. For two of such methods, this dissertation contributed to their adaption for journalism research. Firstly, it helped to make topic modeling more accessible to other journalism scholars by discussing how it can be used in communication studies, and concretely showing how to do it. It also demonstrated an innovative use of search strings to measure various news quality criteria automatically, which may be used for and refined by further research.

**Discussion**

In this dissertation, I used quality criteria for news derived from the electoral model of democracy. It became clear that popular and elite newspapers are not that different in their quality by these standards; only in their online editions the typical difference between the two becomes apparent. I demonstrated that automatic content analysis methods can replace manual analysis for certain steps of the research process and for certain variables, which speeds up data analysis, and as such, allows large amounts of texts to be analyzed. My findings and insights are naturally based on a number of choices. They have certain theoretical and practical implications and suggest a number of avenues for further research, which I will discuss below.

**Democratic Ideals**

As a starting point, this dissertation has used the electoral model of democracy as its normative standard for a 'good' democracy that political coverage should contribute to (Sartori 1987; Strömbäck 2005). Typically, the electoral model holds that voting is the most important form of citizens' democratic influence, that they may participate in the public discourse if they use rational arguments, know what they are talking about, and that political information is only relevant for the political process if it is 'serious', i.e. concerns plans and performances of political actors (Ferree et al. 2002). Even if
we open up the notion of news quality derived from this fairly elitist model somewhat by incorporating audience information-processing skills and expectations, it is still more focused than the more comprehensive participative model or the deliberative model. This has been mentioned before, but it perhaps fitting to emphasize again that I do not mean to argue that the electoral model and its notion of news quality should be preferred over other democratic models in principle. The reason for using this model as a baseline was my premise for this dissertation: given the election campaign investigated here, and given the Austrian political system with its ideal-typical goal that the next Parliament represents the variation of organized opinions in society as closely as possible, I assumed that the electoral model (which shares this goal) provides the appropriate standards for news coverage. For news outside of election campaigns, political news in direct democratic systems, or for news about local politics, where more citizen participation is required, news quality standards derived from other models of democracy could be more suitable. However, my argument that notions of political news quality should incorporate criteria related to increasing news reception still stands, regardless of the democratic model used.

Quality in Elite and Popular News Media

My main empirical finding is that, for most quality indicators measured here, news in elite and popular newspapers does not score that differently, at least in their print editions. More concretely, over the course of the 2013 election campaign in Austria, readers of popular newspapers were not supplied with less news about politics, less diverse news (both in terms of parties and within parties, as focus on leaders), or less complex news, than readers of elite newspapers. This is a good sign for democracy in Austria, and should reduce some of the fears regarding the low attention to politics in popular newspapers by their critics. Of course, these findings only apply to an election campaign, but since the electoral democratic model focuses on elections as the most important time of citizen involvement with politics, news supply during election campaigns is arguably most important.
However, elite and popular media that are highly similar does not mean that both are high in quality. Similar findings by Magin and Stark in non-campaign times were interpreted by these authors as strong tabloidization in Austrian elite media (Magin and Stark 2014). There is no benchmark how much political news, how much diversity and so on is ideal.

The findings presented here do, of course, raise the question to what extent they can be generalized to other contexts. Austria may be an exceptional country with its popularized press market or with its high attention to politics even in its popular press. However, the findings in this dissertation also correspond to an earlier insight by Connell (1998) regarding news discourse in popular and elite media. His findings were based on a case study done in the United Kingdom, which is probably more polarized in terms of popular versus quality than Austria. Connell argued that popularization of news in elite media was actually the incorporation of elements of what he called ‘fabulous reportage’ – a type of news stories typical for tabloids, into news reports of elite newspapers. News reports in tabloids, however, were free from such fabulous elements (Connell 1998). Since most comparative research between elite and popular media compares all news stories rather than particular types (such as news reports) or topics (such as politics) (see e.g. an overview in Reinemann et al. 2012), it could be that the similarity of political news in popular and elite media is not something unique to Austria (or to democratic-corporatist countries with similar press markets as Austria).

It could be something that exists in the ‘typical’ countries for popularization research, such as the United Kingdom or Germany, as well. That is, our perception of political news in popular media could be coloured by the non-political stories in such media (notably tabloids) even though this perception is wrong. A comparison similar to the one made in this dissertation, but for a country with a media market that is more polarized on the popular—elite spectrum could answer the question whether the findings here represent something uniquely Austrian (or democratic-corporatist) or something more universal(-ly Western). Also, I want to emphasize once more that it
would be important to do a similar study outside of election campaigns, because national politics are perhaps less newsworthy for popular newspapers during routine (non-campaign) times (Falasca 2014).

**Quality of Online News**

The small differences between the quality of political coverage in popular and elite news become much larger when we move from the print to the online editions of newspapers (except for news diversity). As mentioned in chapter 3, this could mean that the audience for online popular newspapers has a lower chance to become politically informed than both the audience for online elite newspapers and the audience for print popular newspapers.

This finding needs confirmation from other contexts (routine periods of news, other countries), to see whether it holds there true as well, and also to explain out why this gap in quality exists. Sparks (2003) found a similar gap in British online newspapers. He suggested that the likely reason was that it is difficult for popular news media to make a profit online, since the content that its readers prefer specifically (the entertainment and celebrity sections) are offered by many other (non-news) media in the online environment; media that do not interlace such sections with political or other news. As such, popular news audiences give up online newspapers for non-news sites (*ibid*). In Austria, a similar scenario might hold true as well – here, all newspaper content is freely available online, meaning that the online editions rely on sharing resources with the print version and on what little advertising there may be online. This means that if political news is not very popular with its readers, it will make even less money from advertising. So it might be more attractive business-wise to offer less news about politics. Do newspapers that make only some of their content freely available to subscribers also contain less political news online than in print? It would thus be interesting to gain more insight into the relation between business model, media type, and news quality. The degree of convergence between the online
and the print version of the same paper (García Avilés et al. 2009) is another variable that could be taken into account.

An interesting and important finding of this dissertation is that online editions of newspapers – both popular and elite – are less emotionalized than their print counterparts. So by traditional electoral democratic standards, online editions perform better than print editions on this aspect. As mentioned in chapter 3, what it seems to indicate, however, is an increased use of wire service and news agency material, (descriptive) live blogs and annotated collections of Tweets and photos, and less space for (or total omission of) letters to the editor, editorials or opinion pieces. Further research is needed to prove that this is the actual cause of the lower emotionalization online. But if it indeed is, it would mean that overall political coverage online might actually be poorer than that in print, since its orientation function (Connell 1998), the guidance of its readers towards forming an own opinion through providing opinion and analysis of alternatives, is reduced. Photos, live blogs and press agency material are likely less helpful for doing so than opinion pieces and editorials, even though the latter will probably contain some emotionalized language. To what extent online newspapers contain different ‘bundles’ of news stories (Sparks 2003) compared to their print version, should be investigated more in-depth and beyond an election campaign.

Another interesting question to explore is what exactly the story types that are unique to online editions (notably liveblogs) are about, in terms of political information, emotionalization and other characteristics. And what the consequences for news quality (by electoral democratic standards and in general) could be, both empirically and normatively, of a news offer that includes these new types of news stories, but no longer includes other types that are part of the traditional offer in print newspapers, such as letters to the editor.

Lastly, a limitation regarding the empirical findings for the news quality in Austrian news outlets is the limited range of indicators used. I chose to use these particular indicators because they could be
measured automatically using search strings, or in the case of complexity, because the development of this particular indicator was the goal of the chapter. Search strings have the advantage that they are quick, precise and have high replicability, allowing me to include online outlets, which would have not been possible otherwise, for practical reasons. A limitation of this method is however that it can only measure things that are manifest in the text – words that can be counted. More sophisticated aspects of news quality or popularization, meaning those that need some degree of human interpretation, notably privatization (the focus on the private lives of politicians) or human interest (stories about the experiences of unknown individuals), cannot (yet) be measured with these methods. I thus chose a broad range of outlets over a broad range of indicators of news quality, but further research could focus on a broader range of indicators, using both manual and automatic content analysis.

**New Measures and Methods for Content Analysis**

Aside from theorizing and measuring criteria of news quality, this dissertation contributed to the development of measures and methods to explore news content and quality. I aimed to do so partly in response to a gap in previous studies (namely, the lack of measures for content complexity) and partly in response to challenges for content analysis presented by the current fragmented and infinite media environment. Here, I will discuss the implications of both the findings for and the use of this measure and method, as well as their limitations.

As an indicator of news quality, complexity is a concept that immediately comes to mind when distinguishing between elite and popular news (e.g. Gans 2009) – the former is assumed to be more complex (and according to some, therefore better) than the latter. I argued that complexity has many aspects: news coverage can be complex in its structure as text, in its grammar, in the argumentation used in its content, as well as in the elements of the picture of political reality that it represents. This dissertation showed that for this last dimension of complexity, elite newspapers are not consistently more complex than popular newspapers, for their political coverage. These
findings (combined with the similar findings regarding diversity), are important because they show that over the course of an election campaign, readers of popular newspapers and of elite papers were offered similar possibilities to become well informed about different opinions and issues during the election campaign studied here.

My results show an accurate and comprehensive picture of how complex the content of elite and popular newspapers is. This does not mean, however, that readers actually use this full complexity. But the offer is there—and this is what the norms demand. Still, it would be interesting to further explore the overlap between what people perceive as complex and what the measure presents here as complex, as a way of validating our findings. For example: if our measure says outlet A has content that is more complex than outlet B in a particular period of time, do people then also perceive it as such? Such tests of overlap between what is measured and what is perceived would further add to the usefulness of the new measure proposed here.

However, looking at complexity, represented as networks of actors and issues also has some limitations, especially when seen in the light of the arguments for popularized electoral democratic quality criteria posed in chapter two. In that chapter, following Dahlgren (2000), I distinguished between the thematic and the pragmatic dimension of news. The former contains its content, the latter the presentation of this content. News content complexity measures the complexity of the thematic dimension of news, but does not tell us about the complexity of its pragmatic dimension. Since I argued that the pragmatic dimension should be given more attention in research and more credit in the evaluation of news quality, the findings on complexity shown here are only a first step – whether these findings are ‘good’ for democracy thus also depends on the complexity of news in its pragmatic dimension, meaning its textual, or other presentational, aspects. Are these levels of complexity optimal for the understanding and interest of each outlet’s audience? Although I see the findings regarding complexity thus as a first step rather than a complete picture, they are still important. They show that readers of different
6. Discussion and Conclusion

types of media are presented with news that connects actors and issues in similar ways, rather than showing fragmentation.

The argument regarding validation can also be made for LDA. Here, other studies have already shown some ways to validate outcomes of automatic analysis using human coders (see e.g. Chang et al. 2009; DiMaggio, Nag, and Blei 2013), but there is no generally accepted validation method yet. Validation – showing that the algorithm produces results that reflect the concept that researchers intend to measure – is important for making topic modeling an acceptable research method in the social sciences.

**Avenues for Further Research**

The chapter on LDA was also meant as an introduction to (inductive) automatic content analysis to journalism or political communication scholars not yet familiar with this type of methods, in the hope of increasing interest in the potential of these methods, and in improving their adaption to the goals of communication research. A danger of the adoption of techniques created in one academic field by another is that it is done uncritically, without considering how its findings can be embedded in the concepts of the latter field. By designing this chapter as a replication study rather than a new study and by discussing the interpretation of topics extensively in the chapter, I hoped to avoid this trap. However, understanding the technological side of both the media platforms that we study as well as the (digital) methods that we use to study them has become more challenging with online news, and should be given more attention in communication studies research and curricula as well. Or, alternatively, cooperative projects between computer science and communication studies could be facilitated more. Automatic content analysis allows for the measurement of certain indicators in a matter of hours instead of the weeks or months required for manual content analysis. But not only its speed, but also its precision (measurement on the level of individual words or lemmas) and roots in computational linguistics are promising for combining with language-based content analysis methods, such as discourse analysis. Communication scholars should actively take part in the development
of these methods, instead of leaving them to computational linguists (or political scientists, who seem to be more active in this regard). A good start would be validity measurements of topics in news content based on audience studies mentioned before: Do algorithm and human news consumer agree on what news stories are about?

For the evaluation of news quality, too, I argue that we should concentrate more on audience-or reception-based criteria. The characteristics of news that can be shown to contribute most towards attaining the goals of a particular democratic model should be more important quality standards than a set of ideal content or presentation characteristics that theoretically lead to certain outcomes, or work for one particular audience (elite) that in practice is only a minority. For news quality research, this means that it should in principle get an audience-focus component in addition to content analysis only. This argument still stands regardless of the empirical findings in the third and fourth chapters: these findings in the news ‘as sent’ are only one part of news quality, and their evaluation should depend on how they contribute to or detract from informing citizens and motivating them to vote.

Here, the relevant question is not only which style works best for a particular audience, but also whether all audience groups in a society are covered (a relevant consideration would then also be what exactly an ‘audience group’ is, and whether perhaps certain non-users of political news could be targeted as an audience group). In this way, all of society could be supplied with political coverage that informs and involves as well as possible. This is important to consider in media regulation as well (e.g. public service media should have newscasts aimed at different audience groups, see also Gans 2009; Costera Meijer 2005).

Also, if news quality would include this reception component also in its common perception (what people think of as quality news), and if it could be integrated in journalism education (e.g. how to create news content high in political information, attractive to and understandable by specific audience groups), it could perhaps also
relieve some of the stigma on popular news, both in society as well as among (popular and elite) journalists.

But also for content analysis the evaluation of news quality presented here could and should be extended. Notably, the insight based on the finding that certain news story types are lacking in online editions, and the consideration that this might harm certain functions of political coverage for democracy, implies that the diversity of news story types (news reports, background stories, editorials, service journalism etc.) should also be considered a quality criterion. If we acknowledge that involving people in politics and helping them understand political decisions are also important functions of the media, beyond transmitting information (which, arguably, can be done by live blogs and press agency material just fine if general news quality criteria are adhered to), then we need regard it as a criterion for quality that news, on whatever platform it is distributed, contains different types of news stories that fulfill these different functions.

Aside from this diversity of news story types, the diversity of news sources in online media was also a concern I mentioned earlier in this dissertation: to what extent do online outlets base their coverage on the same (press agency) material? If all outlets were to use the same source, and that source only, this could threaten diversity and thus news quality, since one source is more likely to miss out on a certain viewpoint or issue than a lot of different sources (Napoli 1999). Although the use of the same press agency material by different (online) outlets is increasingly getting attention from researchers (e.g. Paterson 2005; Quandt 2008; Welbers 2015), it is another point that can be considered as an aspect for news quality, and could be included in research as such.

To return to the necessity of different types of news stories, I think that news quality research does not sufficiently distinguish between these different functions of news story types, and is instead rather focused on news reports or background stories. But what is a ‘quality’ editorial, or a ‘quality’ selection of letters to the editor? Notably, for the orientation function of political news (that of helping citizens locate themselves ideologically in the political landscape, helping
them form an opinion by analyzing and discussing certain alternatives), we should find out if some news story types or some discourses are better than others. For example, are reader forums a good replacement of letters to the editor? Again, what is considered ‘quality’ here depends on what works best for achieving the goal of orientational coverage – motivating people to consider different opinions and form their own. Deliberative democratic theory might be helpful for providing some of these criteria, but to see what works well and what not, audience research is again necessary.

And this is where both avenues for further research meet: characteristics of news coverage that have particular effects on audiences are often far more fine-grained and nuanced than the focus on leaders or the use of emotions alone, but might be language-based, and can thus potentially be found in texts using automatic content analysis (and then used in reception studies the traditional way). Particular political discourses (e.g. populism), particular ways of using emotions, aspects of private politicians’ performances or roles of unknown individuals could be more relevant than simply emotionalization or personalization (see e.g. Connell 1998; Fine and White 2002), but quantitative studies on popularization and news quality seem stuck in the usual categories, that are perhaps increasingly missing the point, given developments in online media such as the ones described in the previous paragraph. Meanwhile, automatic content analysis and digital methods in general are making great progress in other fields and outside of academia (Lazer et al. 2009). There is a lot of work to do here for communication scholars, and the quality of news in online media, popular and elite, would be a great case to continue studying in-depth.

References
6. Discussion and Conclusion


Bibliography


The Quality of Political News in a Changing Media Environment


Bibliography


The Quality of Political News in a Changing Media Environment


Tandoc, Edson C. (2014). Journalism is twerking? How web analytics is changing the process of gatekeeping. *New Media & Society*, 16(4), 559-575. Doi: 10.1177/1461444814530541


Welbers, Kasper, Van Atteveldt, Wouter, Kleinnijenhuis, Jan, Ruigrok, Nel, and Schaper, Joep (In press) News selection in the digital era:
professional norms vs. real-time audience metrics. Accepted for publication in *Journalism.*


Account of Funding and Co-Authors

Funding
The research for this dissertation was conducted during my time working for the Media Side of the Austrian National Election Study (AUTNES MediaSide). The Austrian data presented here forms part of this project, and I was a part-time employee of this project while doing the research for and writing my dissertation. AUTNES MediaSide is fully funded by the Austrian Science Fund (FWF); Grant S10908-G11. I received no other funding.

Contributions from Co-Authors

Chapter 2: Evaluating the quality of political news in a popularizing news environment

Carina Jacobi read, summarized and categorized the relevant literature, and wrote and finalized the text.

Wouter van Atteveldt read and commented on the text, and made suggestions for additional sections to include, as well as for revisions in structure.

Klaus Schönbach read and commented on the text, and made suggestions for additional sections to include, revisions in structure, as well as textual revisions.

Chapter 3: Political news content in online and print newspapers

Carina Jacobi did most of the literature review, performed the data analysis and interpreted the data, wrote most of the paper, and finalized the text.

Katharina Kleinen – von Königslöw assisted in writing and revising the chapter in the different versions it went through, by giving comments and shortening the text.
The Quality of Political News in a Changing Media Environment

Nel Ruigrok did part of the literature review, wrote part of the theory section of the paper and helped rewriting the text in the final revision of the chapter.

Chapter 4: Exploring News Content Complexity Using Network Analysis

Carina Jacobi did the literature review, made the decisions regarding which measures to include, assisted in the data analysis, interpreted the results, wrote most of the paper and finalized the text.

Jakob-Moritz Eberl assisted in the data analysis (notably data preparation), wrote part of the ‘Indicators of content complexity from network theory’ and ‘Data and methods’ section, and made suggestions for revising the text.

Katharina Kleinen-von Königslöw wrote part of the ‘Defining content complexity’ section and helped making the argumentation clearer and rewriting the text for the final revision of the chapter.

Stephan Schlögl performed the main part of the data analysis.

Chapter 5: Quantitative Analysis of Large Amounts of Journalistic Texts Using Topic Modeling

Carina Jacobi did the literature review, performed part of the data analysis, interpreted the data, wrote most of the paper, and finalized the text.

Wouter van Atteveldt made suggestions for more literature to include, wrote some of the scripts for the data analysis and helped with their execution, wrote part of the paper (especially in the discussion section) and helped revising the text.

Kasper Welbers gathered the data, performed the analysis for the ‘Granularity’ section and wrote part of this section, created the topic browser mentioned in the text, and helped revising the text.
Summary

The Quality of Political News in a Changing Media Environment

The media environment in Western democracies is changing. The supply of news is shifting from print and electronic media towards (individualized) online media. Online, news audiences are offered an increased choice between different outlets. Also, website technology allows for increased interaction between news producers and news audiences, both visibly (e.g. through comments on news stories) and invisibly (e.g. through click-rates). This changing media environment could be beneficial for news quality, but it could also result in more fragmentation, more possibilities to avoid news that one does not like (intentionally or not), and more popularized news – news with an emphasis on those qualities that make it attractive, interesting and easy to understand for a large audience.

News media are an important means through which people learn about what is going on in the world. The information and opinions these media offer can help people form their own opinion on social issues. Lastly, news media give advice on practical problems, including the micro-level consequences of (nation-wide) policies. Given this broad set of functions that news media have in society, it is important that the coverage in these media is high in quality. During election campaigns, political news forms an important pool of information from which people can tap when deciding for whom to vote. The electoral model of democracy provides a standard of quality for political news that is focused on this information-providing function of news media. According to this standard, political news high in quality is factual and impartial, relevant but diverse, and has a detached and civil style.

This dissertation focused on the question of what the ongoing changes in the media environment mean for the quality of political news, as well as for research on this topic. I firstly discussed if and how we should re-conceptualize political news quality in a
popularizing media environment. I discussed this theoretically, based on findings from previous studies. Secondly, I investigated empirically how different outlets in a particular context (Austria during the 2013 National Election Campaign), performed on various quality aspects derived from this theoretical discussion. Lastly, I looked at how new measures and (automatic content analysis) methods can contribute to exploring news content and quality in a changing media environment, based on the methods and measures used in the three empirical studies in this dissertation.

The dissertation started with a categorization and discussion of earlier findings on the benefits and risks of popular(-ized) news for democracy. By news quality standards derived from the electoral democratic model, popular news is considered to inform citizens insufficiently about political issues. Its emotionalized language and argumentation make it supposedly more difficult for people to form an informed opinion, thus making it low in quality. A second strand of research on popular news instead shows its benefits for democracy: popular news may political interest and knowledge among its audience. Researchers from this strand prefer news quality standards derived from the participatory model for democracy, which emphasizes popular involvement and participation in politics. By these standards, popular news performs much better, according to its supporters.

A problem in many Western democracies is that voter turnout and political interest are falling, especially among the young. Given this problem, and based on the findings regarding the benefits of popular news, I argued that the quality standards derived from the electoral democratic model should be made more lenient towards popular news elements that increase political interest and knowledge among certain audience groups. Political news quality standards should take into account empirical evidence on which elements of news best contribute towards fulfilling the goals of the democratic model from which they are derived– for electoral democracy: citizens who vote, and do so in an informed way. As a result, what ‘quality’ news looks like depends on the news audience: the news style and focus that work for informing the highly educated audience for elite media may
not work for a popular news audience, for example. I argued that amount of political news, relevance, diversity, complexity and detachment are important quality criteria, but that the ideal level of each of these criteria (e.g. how much political news is enough?) should be adapted to the audience.

The two empirical chapters compared the quality of political news between Austrian media outlets during the 2013 Austrian National Election Campaign. Popular newspapers have a large market share in Austria, and Austrian elite newspapers are rather popularized, making this country an interesting, atypical case study.

First, I looked at the differences in online and print editions of six popular and elite newspapers regarding four quality criteria derived from the electoral democratic model: amount of political news, diversity between and within parties (that is, a low focus on political leaders), and low emotionalization. The online media environment is both a challenge and an opportunity for political news in terms of these criteria. Among the challenges are the hasty news cycle and the increased competition compared to the print market. Virtually unlimited space and interactivity online, however, carry great potentials for news high in quality (notably diversity).

Grouping all newspapers together, the online editions seemed to perform better than the print editions: they provide their audiences with fewer, but longer news stories, resulting in more political news overall. Online editions also showed higher diversity between parties and less emotionalization, but a stronger focus on leaders.

However, this picture changed when I looked at popular newspapers and elite newspapers separately. In print, popular newspapers performed even slightly better than elite papers due to the great amount of political news they offer. However, they performed worse in their online editions: they provide less political news online than in print, and are more focused on political leaders. Furthermore, there is less space for opinion pieces in their online edition: editorials, letters to the editors and other opinion pieces were replaced by a readers’ forum in two popular outlets. In the other outlets, editorials and opinion pieces were part of the online edition. Overall, differences
between popular and elite newspapers were larger online than in print, to the disadvantage of popular newspapers. I argued that the lack of opinion section in two of the three popular online papers harms the orientation function of media – helping the audience form an opinion. The larger gap in quality between elite and popular outlets online may reinforce the gap between a better informed elite audience and a lesser-informed popular news audience, especially because young people choose the Internet over other news channels.

Secondly, I investigated one aspect of news quality, content complexity, in-depth. While theoretically relevant for the discussion about news quality and popular news (lower complexity can increase understandability, but may also result in less informative political news), content complexity is rarely included in empirical news quality research. To the complexity of news content, displayed as networks of actors and issues, I used three dimensions of complexity derived from network analysis (graph theory), and their corresponding measures. Content complexity informs us how complexly news outlets depict the political arena. Interconnectedness measures how many connections are made between actors and issues. The more actors engaged in a debate or the more issues are connected to each other, the more complex the political arena. Centralization measures how strongly the political arena as reflected in the news is focused on particular actors or issues (thus reducing complexity). The more different actors speak or are being spoken to, or the more different issues are being connected, the more complex the political arena. Heterogeneity measures how strongly this political arena focuses on particular actor groups to speak to or act towards others, or on particular ideological issue poles that are contextualized with other ideological poles (again reducing complexity). The more actor groups (such as coalition parties, citizens, experts or companies) speak or act towards various other actor groups, or the more issues important to a specific ideology (such as conservative issues, economic right issues, or green issues) are connected to issues important to various other ideologies, the higher complexity.

Content complexity was not higher in elite newspapers than in popular newspapers across the three dimensions. Differences
between outlets were small, although one (free) popular paper stood out for its low interconnectedness. However, the two types of papers let different actor groups speak or act in the news, and emphasize issues from different ideological poles.

Overall, we cannot simply say that political news in Austrian popular papers showed lower quality than in elite papers during the 2013 National Election Campaign, at least in their print editions. The view of popular news as apolitical, sensationalized and simplified does not hold true in this context. A similar study for a country with a media market that is more polarized on the popular—elite spectrum could answer the question whether the findings here represent something uniquely Austrian or something more universal(-ly Western) for political news during election campaigns.

The aforementioned two studies also contributed towards creating new measures and methods suitable for the study of news content and quality in a changing media environment.

Following my argument that different outlets should provide political news content suitable for their respective audience groups, they should also provide news with different levels of (content) complexity. The measure for content complexity introduced here can be a first step towards measuring the extent to which media fulfill this need. Especially the interconnectedness dimension is conceptually related to difficulty: more complex news has more explicit connections between its elements, and can thus be perceived as denser in information, which is presumably harder to understand. Further research is needed to validate this.

The comparative study on online and print media used the automatic content analysis method text mining (measuring indicators using search strings) to measure aspects of political news quality. This method is faster and more reliable than manual content analysis, but limited in what it can measure in terms of aspects of news quality. I therefore suggested that further research should find ways to combine automatic and manual content analysis in an optimal way, in order to benefit from the advantages of both.
Finally, the last study fully focused on a relatively new research method: the topic model Latent Dirichlet Allocation (LDA), an algorithm that automatically clusters words into topics based on patterns of co-occurrence in texts. This method has the potential of automatically coding the topics of news texts, which would be very useful in a media environment where news content is abundant and media use is fragmented. Topics might say something about political news quality, given that the electoral democratic model has strict relevance criteria. However, topics are only useful if they overlap with theoretical concepts of interest, in particular issues or frames. To investigate this overlap, I compared the results of an LDA analysis of the change in news topics over time with the results of an existing study: that of Gamson and Modigliani (1989) regarding the changing framing of issues relating to nuclear technology in American news coverage since the Second World War.

Although there was some overlap, the topics found through LDA topic modeling did not approach ‘frames’ in the sense of coherent interpretative packages. They also did not represent valence or sentiment. Topics instead represent various kinds of regularities in language use across texts: issues on different levels of abstraction, events, or writing style. As such, the outcome of an LDA analysis is a first step rather than a complete analysis in itself: it can help researchers find relevant texts to include in a sample, help with the creation of codebooks for manual content analysis, or give a quick first impression of the contents of a very large set of news texts. Further research could investigate if more advanced forms of LDA, which allow researchers to steer the clustering of words in certain ways, could automate a larger share of the content analysis process.
Samenvatting (Dutch-language Summary)

De kwaliteit van politiek nieuws in een veranderend medialandschap

Het medialandschap in westerse democratieën is in verandering. Het aanbod van nieuws verplaatst zich meer en meer van televisie en papieren kranten naar (geindividualiseerde) online media. Online heeft het nieuwspubliek meer keuze tussen verschillende media en is er meer interactie tussen nieuwsconsumenten en –producenten, zowel zichtbaar (middels commentaar onder artikelen) als ook onzichtbaar (het volgen van ‘clickrates’ van lezers). Dit veranderende medialandschap kan een positieve uitwerking hebben op nieuwskwaliteit, maar het kan ook resulteren in een meer gefragmenteerde mediamarkt, meer mogelijkheden om nieuws te ontwijken dat men niet bevalt (met opzet of onbedoeld) en meer gepopulariseerd nieuws – nieuws waarvan de prioriteit ligt op de aantrekkelijkheid en begrijpelijkheid voor een groot publiek.

Via het nieuws leren mensen over de gebeurtenissen in de wereld. Nieuwsmedia helpen hen niet alleen om geinformeerd te blijven maar ook om een mening te vormen over sociale problemen en geven informatie over allerlei praktische problemen, waaronder de mogelijke consequenties van landelijk beleid voor hun dagelijks leven. Kwaliteitsjournalistiek is belangrijk gezien deze brede reeks van functies van nieuwsmedia in de samenleving. Een van de momenten dat deze media vooral van belang zijn, is tijdens verkiezingscampagnes. Dan vormt nieuws in massamedia een belangrijke bron van informatie voor kiezers, mede op basis waarvan zij besluiten op wie zij zullen stemmen. Vanuit het electorale democratiemodel is een aantal criteria opgesteld die de kwaliteit van (politieke) nieuwsvoorziening kunnen bepalen, die vooral betrekking hebben op deze informatieverlenende functie van nieuwsmedia. Volgens deze standaards is kwaliteitsnieuws feitelijk en onpartijdig.
relevant doch divers; qua stijl is kwaliteitsnieuws afstandelijk en beleefd.

Dit proefschrift ging over de vraag wat de doorlopende veranderingen in het medialandschap betekenen voor de kwaliteit van politiek nieuws en voor onderzoek naar dit onderwerp. Aan de orde kwam ten eerste de vraag of en hoe we de kwaliteit van politiek nieuws een nieuwe betekenis moeten geven in een populariserend medialandschap. Hier ging ik theoretisch op in, op basis van bevindingen van eerder onderzoek. Ten tweede bekeek ik empirisch hoe nieuwsmedia in een specifieke context (Oostenrijk tijdens de Nationale Verkiezingscampagne van 2013) presteerden in termen van kwaliteit. Tot slot ging ik in op de vraag hoe nieuwe onderzoeksmatten en – methoden (met name op het gebied van automatische inhoudsanalyse) van nut kunnen zijn bij dergelijk onderzoek, op basis van de maten en methoden die ik gebruikte in de drie empirische studies in dit proefschrift.

Het proefschrift begon met een categorisering van eerdere bevindingen op gebied van de voor- en nadelen van populair nieuws voor een democratie. Vergeleken met de idealen van het electorale democratiemodel informeert populair nieuws burgers onvoldoende over politiek, en bemoeilijkt de focus van populair nieuws op emoties van populair nieuws het vormen van een onderbouwde mening over politieke onderwerpen. Een tweede groep studies toont juist de voordelen van populair nieuws voor een democratie: het verhoogt de politieke interesse en kennis onder het publiek voor dit type nieuws. Onderzoekers uit deze groep verkiezen kwaliteitsstandaards afgeleid uit het participatieve democratiemodel boven het electorale. Participatieve model legt de nadruk op politieke interesse en – participatie onder burgers; nieuws dat hiertoe bijdraagt is hoog in kwaliteit. Bij deze standaard doet populair nieuws het beter, aldus de voorstanders van populair nieuws.

Een probleem in veel westerse democratieën is dat de opkomst bij verkiezingen terugloopt en dat politieke interesse daalt, vooral onder jongeren. Gezien dit probleem en gezien de eerder genoemde bevindingen ten aanzien van de democratische bijdrage van populair
Samenvatting

nieuws, stelde ik dat de kwaliteitsstandaards voor politiek nieuws versoepeld moeten worden. Nieuwselementen die politieke interesse en kennis onder een bepaald publiek verhogen zouden ook als kenmerken van kwaliteit moeten worden erkend. Oftewel, kwaliteitsstandaards zouden empirisch bewijs in acht moeten nemen dat aantoont welke nieuwselementen het beste bijdragen aan het bereiken van de doelen van het democratisch model waarvan deze standaards zijn afgeleid. Voor het electorale democratiemodel is zo’n doel bijvoorbeeld dat burgers geïnformeerd zijn over politiek, en dat zij stemmen bij verkiezingen. ‘Kwaliteitsnieuws’ zal er hierdoor anders uitzien, afhankelijk van het publiek: bijvoorbeeld, de nieuwsstijl en focus die goed werken om een elitepubliek over politiek te informeren werken waarschijnlijk minder goed voor het publiek van populaire nieuws. Ik stelde vervolgens dat de hoeveelheid politiek nieuws, relevantie, diversiteit, complexiteit en (emotionele) afstandelijkheid belangrijke kwaliteitscriteria zijn voor politiek nieuws, maar dat het afhankelijk is van het publiek hoe sterk elk van deze criteria aanwezig zou moeten zijn.

De twee empirische hoofdstukken vergelijken aspecten van kwaliteit in politiek nieuws tussen verschillende Oostenrijkse media gedurende de Oostenrijkse Nationale Verkiezingscampagne van 2013. Het hoge marktaandeel van populaire kranten en de relatief sterke popularisering van kwaliteitskranten maken Oostenrijk een interessante, atypische casus.

Eerst bekeek ik de verschillen tussen online- en printedities van zes populaire en elite (‘kwaliteits’) kranten voor vier kwaliteitscriteria afgeleid uit het electorale democratiemodel: hoeveelheid politiek nieuws, diversiteit tussen en binnen partijen (dat wil zeggen, geen sterke focus op de lijsttrekker), en lage emotionalisering. Het online medialandschap is zowel een uitdaging als ook een kans voor het maken en publiceren van kwaliteitsnieuws in termen van deze criteria. De snelle nieuws cyclus en hoge concurrentie tussen media vormen een uitdaging. De vrijwel ongelimiteerde ruimte die websites bieden als ook de mogelijkheden voor interactie met het nieuwpubliek kunnen helpen bij het maken van nieuws hoog in kwaliteit (met name op gebied van diversiteit).
Over het algemeen leken de online edities het beter te doen dan de printedities: online edities bieden minder, maar langere artikelen, met als resultaat een grotere hoeveelheid politiek nieuws in totaal. Daarnaast was ook de diversiteit tussen partijen hoger en emotionalisering lager online. De focus op lijsttrekkers was echter sterker.

Dit beeld veranderde enigszins wanneer we populaire en elitekranten apart bekijken. Populaire kranten deden het zelfs iets beter dan elitekranten in hun printeditie door de grote hoeveelheid politiek nieuws in die zij hun lezers bieden. In de online editie daarentegen scoorden ze minder goed op de verschillende criteria: ze bieden minder politiek nieuws online dan in de printeditie en focussen daarbij meer op lijsttrekkers. Ook is er minder ruimte voor opinie online: columns, lezersbrieven en opiniestukken waren in de online versie van twee populaire kranten vervangen door een lezersforum. In de overige media waren opiniestukken en columns wel opgenomen in de online versie. Over het geheel gezien waren de verschillen in kwaliteit tussen populaire en elitekranten groter online dan in print, waarbij populaire kranten in het nadeel zijn. Ik stelde dat het ontbreken van opiniesecties in de onlineversie van twee van de drie populaire kranten ten koste gaat van de oriëntatiefunctie van journalistiek, die lezers helpt een mening te vormen, in deze online versies. Het grotere verschil in nieuwskwaliteit tussen elite en populair online kan de kloof tussen een goed geïnformeerd elitepubliek en een minder goed geïnformeerd publiek voor populair nieuws vergroten, ook aangezien jongeren online nieuws verkiezen boven andere mediakanalen.

In het tweede empirische hoofdstuk ging ik in op inhoudscomplexiteit. Dit is theoretisch relevant voor de discussie over nieuwskwaliteit en populair nieuws: lagere complexiteit kan nieuws gemakkelijker maken om te begrijpen, maar kan ook resulteren in minder informatief nieuws. Inhoudscomplexiteit is echter nauwelijks empirisch onderzocht. Netwerkanalyse (graph theory) gebruikt verschillende maten om verschillende dimensies van de complexiteit van netwerkgrafieken te meten. Ik gebruikte drie van deze maten om drie dimensies van complexiteit te meten die zinvol zijn in de context
van nieuwsinhoud. Opgeteld laat inhoudscomplexiteit zien hoe complex media de politieke arena weergeven. *Interconnectedness* meet hoeveel verbanden worden gelegd tussen actoren of onderwerpen. Hoe *meer* er actoren met elkaar in debat zijn of hoe *meer* er onderwerpen met elkaar in verband worden gebracht, des te complexer het nieuws. Centralisering meet hoe sterk deze politieke arena draait om bepaalde actoren of onderwerpen (wat complexiteit verlaagt). Hoe meer *verschillende* actoren aan het woord zijn of aangesproken worden, of hoe meer er *verschillende* onderwerpen met elkaar in verband worden gebracht, des te hoger complexiteit. Heterogeniteit meet hoe sterk de politieke arena focust op bepaalde groepen actoren of op issues uit bepaalde ideologische categorieën (wat wederom complexiteit verlaagt) Hoe meer *groepen actoren* (bijvoorbeeld coalitiepartijen, burgers, experts en het bedrijfsleven) spreken of handelen tegenover *verschillende andere groepen*, en hoe meer *issues uit een bepaalde ideologie* (bijvoorbeeld conservatief, economisch rechts, of groen) in verband worden gebracht met issues uit *verschillende andere ideologieën*, des te hoger complexiteit.

Inhoudscomplexiteit was niet consistent hoger in elitekranten dan in populaire kranten. Verschillen tussen kranten waren klein, hoewel één (gratis) populaire krant opviel door lage interconnectedness. Echter, de twee mediatypen legden de nadruk op verschillende actorengroepen en issuecategorieën.

Over het geheel gezien was politiek nieuws in de printversie van Oostenrijkse populaire kranten niet lager van kwaliteit dan in elitekranten, tijdens de Nationale Verkiezingscampagne van 2013. Het beeld van populair nieuws als apolitiek, sensationeel en simpel klopte niet in deze context. Soortgelijk onderzoek in een land met een sterker gepolariseerde mediamarkt (in termen van elite versus populair) kan de vraag beantwoorden of het hier om een uniek Oostenrijks fenomeen gaat of om een meer universeel (Westers) fenomeen voor politiek nieuws tijdens verkiezingscampagnes.

De twee hierboven beschreven studies droegen ook bij aan het creëren van nieuwe onderzoeksmethoden en -maten die passen bij
het bestuderen van nieuwsinhoud en -kwaliteit in een veranderend medialandschap, het laatste onderzoeksdoel van mijn dissertatie.

Uit mijn theoretische argument dat politiek nieuws aangepast moet worden aan de publiekgroep volgt dat media(-typen) nieuws met verschillende niveaus van complexiteit aan moeten bieden. De hier geïntroduceerde maat voor inhoudscomplexiteit vormde een eerste stap ter beantwoording van de vraag in hoeverre media aan dit ideaal voldoen. Vooral de interconnectedness-dimensie is in theorie gerelateerd aan moeilijkheid: complexer nieuws volgens deze dimensie legt meer connecties tussen verschillende elementen, is daarmee dichter in informatie en waarschijnlijk moeilijker te begrijpen. Verder (publieks-)onderzoek is nodig om te laten zien of dit inderdaad het geval is.

De vergelijkende studie naar online en printmedia gebruikte de automatische inhoudsanalysemethode text mining (indicatoren meten met gebruik van zoektermen) om aspecten van politieke nieuws kwaliteit te meten. Deze methode was sneller en betrouwbaarder dan handmatige inhoudsanalyse, maar had zijn grenzen met betrekking tot de kwaliteitsaspecten die ermee gemeten kunnen worden. Verder onderzoek zou zich kunnen richten op het combineren van automatische en manuele inhoudsanalyse, zodat de voordelen van beide methodieken optimaal tot hun recht kunnen komen.

Tot slot richtte ik mij in de laatste studie op een relatief nieuwe onderzoeksmethode: het topic model Latent Dirichlet Allocation (LDA). Dit algoritme clustert woorden automatisch in topics, gebaseerd op woorden die vaak samen voorkomen in teksten. Deze methode zou deze topics mogelijk geheel automatisch kunnen coderen, wat nuttig is in een gefragmenteerd medialandschap met een overvloed aan nieuws. De topics in het nieuws kunnen bovendien relevant zijn voor de discussie rond de kwaliteit van (politiek) nieuws, aangezien sommige onderwerpen meer met hoge kwaliteit worden geassocieerd (bijvoorbeeld de financiële plannen van de verschillende politieke partijen) dan andere (bijvoorbeeld het privéleven van politici). Echter, topics zijn alleen nuttig mits ze
overlappen met concepten die theoretisch gezien relevant zijn, zoals onderwerpen of frames. Om deze overlap te onderzoeken vergeleek ik de resultaten van een LDA-analyse naar de verandering in nieuws topics over tijd met de resultaten van een bestaande studie. Deze studie was die van Gamson en Modigliani (1989) naar de verandering in framing van nieuws over nucleaire technologie in Amerikaans nieuws sinds de Tweede Wereldoorlog.

Hoewel er enige overlap was, zijn topics niet te behandelen als frames in de zin van coherente interpretaties van een bepaald onderwerp. Topics bevatten geen waarde of sentiment. Zij representeren allerlei regelmatigheden in taalgebruik binnen een groep teksten: onderwerpen op verschillende niveaus van abstractie, gebeurtenissen, of schrijfstijl. De resultaten van een LDA-analyse zijn dus beter te behandelen als eerste stap dan als zelfstandige analyse. Zij kunnen onderzoekers helpen om relevante teksten te vinden voor strategische steekproeven, om codeboeken te ontwerpen voor handmatige inhoudsanalyse, of om een snelle eerste indruk te geven van de inhoud van een grote set teksten. Verder onderzoek zou zich kunnen richten op de vraag of geavanceerder vormen van LDA, waarmee het clusteren van woorden meer gestuurd kan worden, toepasbaar kunnen zijn om een groter deel van het inhoudsanalyse-proces te automatiseren.
Acknowledgements

I want to thank my supervisor and co-supervisors, the doctorate committee, my co-authors, and all colleagues, friends and family who contributed to this dissertation directly or indirectly. A special mention goes to Nel Ruigrok and Wouter van Atteveldt, to my parents, and to Mike Phillips: thank you for all your help and support.