Nick Anstead and Ben O'Loughlin
Social media analysis and public opinion: the 2010 UK General Election

Article (Accepted version)
(Refereed)

http://eprints.lse.ac.uk

Social Media Analysis and Public Opinion: The 2010 UK General Election

Abstract

Social media monitoring in politics can be understood by situating it in theories of public opinion. The multi-method study we present here indicates how social media monitoring allow for analysis of social dynamics through which opinions form and shift. Analysis of media coverage from the 2010 UK General Election demonstrates that social media are now being equated with public opinion by political journalists. Building on this, we use interviews with pollsters, social media researchers and journalists to examine the perceived link between social media and public opinion. In light of competing understandings these interviews reveal, we argue for a broadening of the definition of public opinion to include its social dimension.

Social Media Analysis and Public Opinion: The UK General Election 2010

<u>**Introduction**</u>

Recent years have seen a great deal of academic interest in the possibility of using social media to measure public opinion. Broadly this research has looked to achieve two goals. First, scholars have attempted to use social media data to predict election results (Franch, 2013; Jensen & Anstead, 2013; O'Connor, Balasubramanyan, Routledge, & Smith, 2010; Tumasjan, Sprenger, Sandner, & Welpe, 2010; Tumasjan, Sprenger, Sandner, & Welpe, 2011). A second strand of research has attempted to measure the public's evolving response to stimuli, examining both short term events such as televised political debates (Elmer, 2013; Shamma, Kennedy, & Churchill, 2009) and long term events such as economic downturns (Gonzalez-Bailon, Banchs, & Kaltenbrunner, 2010).

Such studies have been criticised for a number of reasons. It has been argued that social media research of public opinion involves arbitrary methodological decisions, such as which political parties to include in the study (Jungherr, Jürgens, & Schoen, 2012). This problem is made more acute, critics argue, because research is often only conducted after the event studied. As a result, various methods can be tested until one appears to work. However, this greatly limits the general applicability of any "successful" approach (Gayo-Avello, 2011). Additionally, it has also been noted that the user-base of social media is not representative of the voting population and, as yet, no methods exist to correct this (Gayo-Avello, 2013; P. T. Metaxas, E. Mustafaraj, & D. Gayo-Avello, 2011).

While these debates have been useful and vigorous, they have almost exclusively focused on questions of method. This article seeks to correct that oversight, and address the relationship between social media and public opinion through the prism of public opinion theory. In particular, while there are clearly shortcomings in social media-based public opinion research

(especially due to the non-representative user-base), the extent to which these are viewed as insurmountable problems largely depends on the definition of public opinion being employed.

It has long been noted that public opinion is a very slippery concept to define. As a result of this, there are numerous definitions of public opinion in circulation (Childs, 1939; Stromback, 2012). The orthodox contemporary view of public opinion, most clearly expressed by early pollsters such as George Gallup (1939) and Floyd Henry Allport (1937), is that it is nothing more than the cumulative preferences of individual citizens. With this view, public opinion is seen as something that exists and which pollsters aim to discover. More critical definitions have argued against this position, claiming that public opinion research manufactures public opinion rather than simply discovering it (Bourdieu, 1979). Despite these critical voices, there can be no doubt that the Gallupian paradigm of public opinion has continued dominate the journalistic, political and general imagination (Herbst, 1993; Moon, 1999).

A different way to understand public opinion is to approach the concept as multi-faceted and historically contingent. While opinion polling is central to the contemporary definition of public opinion, to suggest that this understanding is permanent would be decidedly ahistorical and neglect a number of earlier definitions and debates surrounding public opinion reaching back to the nineteenth century (Splichal, 2012). Susan Herbst, for example, uses the term *public opinion infrastructure* to describe these distinctive eras of public opinion definition and measurement (Herbst, 2001). The contemporary era of public opinion research is, in this view, seen as the product of a specific infrastructure of opinion polling methods, mass media and modernity.

It follows that new methods used in public opinion research, such as social media analysis, necessitate a reinterpretation of public opinion. Using the 2010 UK General Election as a case study, this article employs a three-stage grounded theory approach to better understand social media-based public opinion research and to offer such a reinterpretation of public opinion. The first stage of this process is a qualitative analysis of media coverage referencing social media during the 2010 election. This analysis suggests that social media was employed in a number of ways to reflect public opinion. Most interestingly, we find references to large-scale semantic analysis of social media data in media coverage (a process we refer to as "semantic polling"). Building on this, the second stage of the research explores this phenomenon further by drawing on interviews with experts in social media and public opinion research. These allow us to explore how different types of researchers conceive the idea of public opinion and the tensions that exist between their conceptions.

Finally, we draw on this dataset to rethink the idea of public opinion. We argue that semantic polling requires a broader definition of public opinion than found in the dominant opinion polling paradigm.. In order to construct such a definition, we argue that there is value in looking to older theories, going back to the nineteenth century political theorist James Bryce and the mid-twentieth century sociologist Herbert Blumer. Their theories differ from the dominant contemporary public opinion paradigm, and are much better able to take account of the nature of social media, which can be understood as an "organ of public opinion" (Bryce, 1888). Understood through this definition, social media public opinion research may not offer a representation of the entire public but it has other virtues. It enables studies that focus on social interaction and conversation rather than simple preferences, and introduces a strongly temporal dimension to public opinion research. Ultimately, if accepted, this conception leaves open the possibility of social media-based methods co-existing with more traditional polling-based techniques, rather than being seen as inferior or bogus.

Motivations, methods and key research objectives

Our interest in social media and public opinion started when [Author 2] was invited to work with a firm carrying out an exploratory study of public responses to the 2010 Prime Ministerial debates on Twitter as part of a larger government funded research project on social media monitoring.[i] To further our understanding of this relationship, we employed a number of methods between April 2010 and June 2012. This involved both a qualitative thematic analysis of media coverage during the 2010 election, and interviews and workshops with experts in the field.

The purpose of the media analysis was to assess how social media data were being used to inform discussion of public opinion in political coverage. In order to examine this, we constructed a dataset of media coverage referencing Twitter during the course of the statutory campaign period of 6th April – 6th May 2010. While Twitter is clearly not synonymous with social media, at this moment in time it provokes the greatest interest as a possible tool for measuring public opinion both in academia and the corporate sector. There are two reasons for this. First, the micro-blog format (for which Twitter is the de-facto standard) lends itself to gathering large datasets in real-time as major events occur. Additionally, the relative openness of Twitter, with the vast majority of posts being publicly available, means data are readily accessible (Gayo-Avello, 2013: 650).

Our search for Twitter-related news content generated 287 items:

- Newspaper articles referencing both Twitter and the general election. This element of the dataset was gathered by using a Lexis Nexis search applied to all national British newspapers (n=227).[ii]
- A television sample, made up of episodes of BBC Newsnight and the ITN Evening News from the statutory campaign period (n=41).

- Additional documents available online relevant to this research, including articles and blog entries published by mainstream media websites, as well as press releases and documents issued by social media analysis firms. These were gathered by monitoring published media and the outputs of social media monitoring firms. These data adds additional insights that are sometimes not found in other media discussions (n=19).

While these data could have been examined using a quantitative media content analysis approach, we opted for a thematic analysis. The reason for this, as earlier research on public opinion has noted (Bauman & Herbst, 1994: 135), is that the amorphous nature of the subject makes formal coding very disparate and thus unsuitable for exploratory research of the kind being done here. In contrast thematic analysis is a powerful tool for theory building from empirical data (for an overview of the thematic analysis method see Boyatzis, 1998 and Guest et al, 2011).

We also conducted a set of elite semi-structured interviews, carried out by research assistants working on the project.[iii] The sample of twenty interviewees was guided in the first instance by our thematic media analysis and then used a snowball technique to broaden the sample. The interviewees fell into five (on occasions overlapping) cohorts: data analysts, who worked for the firms that were actually mining and analysing social media data; the journalists who were reporting these findings; politicians and political consultants; pollsters, who were experts in using more traditional means of measuring public opinion; and electoral regulators, with a responsibility for ensuring the integrity of the electoral process. The sample was chosen with the aim of obtaining a rounded view of the problems and potentials of social media analysis. In order to ensure they could talk candidly, interviewees were assured that they would not be identified (an anonymous list of interviewees is included as appendix one).[iv] Finally, we convened two workshops in a major city in Europe to discuss this

phenomenon which bought together representatives of analytics firms, pollsters, news editors and non-government organizations, giving them a space to allow them to debate conceptual, practical and ethical challenges raised by social media analysis.

This mixed-method approach yielded a rich corpus for the analysis of conceptual and normative understandings of public opinion held by members of "the industry" and how traditional pollsters and journalists felt their roles were challenged by new technologies and analytical methods. This allowed us to deploy a heuristic analysis consistent with the precepts of grounded theory to analyse key themes in the dataset (Glaser & Strauss, 1968; Urquhart, 2013), to reflect on theory and to suggest theoretical development relating to social media analysis and public opinion.

Twitter and news media presentation of public opinion in 2010

How did journalists connect social media and public opinion during the campaign? Our thematic analysis of the 227 newspaper articles as well as 41 television news broadcasts from the 2010 UK election period suggests that references to Twitter were ever-present in the four week campaign. The vast majority of this coverage focused on a few topics, notably politicians who were embarrassed by ill-advised Tweets or Prime Minister Gordon Brown's wife Sarah Brown who had a million followers on the micro-blogging site. However, our combined dataset of newspaper articles, television broadcasts and other election texts produced by media organisations and research firms contains 47 attempts to employ social media (in a number of ways, discussed below) as evidence of public opinion during the campaign.[v]

Previous research has found that public opinion is presented in various ways in media coverage of politics (Lewis, 2001; Lewis, Wahl-Jorgensen, & Inthorn, 2004). This variety of

approaches continues when social media data are used. It was possible to identify three

distinct ways in which social media was employed as evidence. Furthermore, the different

types of data used reflected quite different characterizations of public opinion. These are

shown in Table 1.

[Table 1 about here]

First, and in a manner noted in previous research examining the use of selected Tweets in

mainstream American media (Wallsten, 2011), reporters selectively quoted individual users

of social media to create anecdotal evidence of the public's reaction. Essentially, this is a

form of electronic vox pop. This type of content therefore reflects the long-established

journalistic practice of taking the opinions of individuals to reflect wider strands of opinion in

society (Larson, 1999; Robinson, 2012: 10). Examples of this approach appeared in both

national newspapers and broadcast television, notably in the aftermath of the televised

debates. In the print media, following the first debate, a number of commentators quoted

individual status updates to note that some members of the public were mocking the process

and the three party leaders (Sawyer, 2010; Thorpe, 2010). BBC Newsnight, a "broadsheet"-

style current affairs magazine programme, ran segments in all three of its post-debate shows.

Fronted by reporter Justin Rowlett and employing the website Twitterfall, these segments ran

for up to ten minutes, drawing on social media data, including comments made on Twitter,

Facebook and the BBC's own web forums (Twitterfall, 2012). This information was

deployed largely in an illustrative fashion to suggest the public's reaction to the debates.

However, the positioning within the programme of such reporting of anecdotal social media

commentary was notable. While traditional opinion polls headlined the programme and were

treated with reverence, social media data are normally discussed towards the end of the

broadcast, alongside other, less established methods for gauging public response, such as "the

worm" (a graphical display showing a focus group's response to specific portions of the debate). This section of the programme was also frequently book-ended by lighter discussion of the debates, involving, for example, body language experts. Hence, this running order is indicative of the assumed sovereignty and greater significance of traditional polling.

The second strand of reporting employing social media as evidence for public opinion was based on what we have termed raw quantitative statements. In practice, this took two forms. First, it meant citing the number of tweets or tweeters commenting on a specific topic or event. Interestingly, similar data could be interpreted in many different (sometimes diametrically opposing) ways. For example, on some occasions the numbers were quoted as evidence of positive engagement with the political process, especially among young people:

> [T]housands of people, and especially first-time voters, were watching them [the debates] on two screens: the TV screen and their mobile phone or computer, which they used to monitor and respond on Twitter and Facebook… in the third debate there were 154,342 tweets tagged "#leadersdebate", coming at 26.77 tweets a second, spread among 33,095 people. Though that surely doesn't include many hundreds - perhaps thousands - of others who didn't use the tag (Arthur, 2010).

However, the very same data were also be used to suggest growing boredom and disengagement among the population with the election campaign:

[Tweetminster] reporting this week that were 54,000 tweets, about 27 tweets a second.[vi] Interestingly though, only about 33,000 actual twitterers. So, a small number of twitterers making a lot of tweets out there (Rowlett, 2010).

A second raw quantitative use of social media data involved reference to so-called trending topics. A number of sites monitor popular topics of conversation, and the words and phrases appearing on Twitter in order to publish lists of those that are being discussed by lots of people. The fluctuation of trending topics, and particularly the rapid rise of hashtags (a short code inserted into a message to make it searchable), were widely reported during the election campaign. Some examples that achieved coverage were the result of party mobilisation, such as the Labour Party's #welovethenhs (Hinsliff, 2010). However, by far the most commented on hashtag during the campaign was #NickCleggsFault, an ironic response to a battery of negative stories launched by Conservative supporting newspapers following the Liberal Democrat leader's strong showing in the first debate, when:

> [A] Twitter campaign made a mockery of the crude smears. The "nickcleggsfault" tag became the most popular in Britain within hours. Thousands of fun messages blaming him for all the world's problems left the barely disguised Tory attack against the Liberal Democrat looking daft (Roberts, 2010).[vii]

We found statements linking both the number of updates around events or trending topics on social media to share two characteristics. First, the data were used in a very crude form, often simply involving the citing of numbers. Second, the inferences that were drawn from these

data tended to cast public opinion or a sub-set of public opinion as having a single, shared preference. This might be related to their level of engagement or apathy, or possibly indicate anger towards the political classes or the press, as in the #NickCleggsFault example. As such, this type of reporting is analogous to past coverage of electoral campaigns in the UK that might have referenced electoral turnout or incidences of political mobilization by groups with a shared viewpoint.

The third use of social media data are what we term *semantic polling*. The semantic polling process has three elements. First, it involves the harvesting of large datasets from social media services online (in 2010, this meant datasets almost exclusively from Twitter). Second, those datasets are analysed using computer-based natural language processing techniques to attribute some kind of numeric indicator of sentiment to it, most often relating to the number of positive to negative comments about specific politicians, parties or policies. Finally, this information is put into a format, numerical or graphical, suitable for public dissemination.

It could be argued that this method is not really akin to polling, at least as we traditionally understand it. After all, the sample is non-representative of the voting population (an issue discussed further below), while no citizens have actually had their opinions directly solicited.[viii] However we would argue that the term polling is appropriate for two reasons. First, semantic polling presents public opinion as being heterogeneous, and aims to measure the differing attitudes and reactions that exist among citizens. Second, semantic polling can produce quantitative public opinion information that can be discussed and presented either numerically or graphically in a very similar manner to traditional opinion polls. As such – and while the methods of data collection might be very different – the outputs of semantic polling are of a similar genre to traditional opinion polls.

A number of private firms specializing in these techniques appeared in election-time media coverage, which featured data produced by Lexalyticals, Linguamatics, Meltwater Buzz, Semiocast and Tweetminster. The findings of these firms were reported by the BBC's technology correspondent Rory Cellan-Jones on his blog on the BBC website, dot.Rory (Cellan-Jones, 2010b). The magazine *PR Week* commissioned a weekly online sentiment tracker (Sudhaman, 2010). Many of the social media consultancies also published their own blogs and press releases.

This media analysis reveals competing understandings of public opinion and the ways journalists report on social media in order to make claims about it. In the next section we turn to interviews conducted with various sections of "the industry" and find such differences are grounded both in the value attributed to different methodologies and broader differences of goals and interests. This analysis provides a foundation for our attempt to re-theorise public opinion in a way that explains how both traditional and sematic polling can complement each other to provide have a more holistic and richer understanding of public opinion.

Semantic polling considered

While useful, media analysis can only take us so far in terms of understanding semantic polling. Our interview data sheds more light on this emerging technique. In particular, the analysis of the semi-structured interviews with social media analysts, traditional pollsters, journalists, political campaign managers and election regulators revealed three interesting findings: the differing techniques being employed by those under-taking semantic analysis of social media data; the relationship between the traditional news media and social media analysis of public opinion; and the very different conceptions of public opinion that existed within our sample of interviewees.

Methods of semantic polling

Reflective of the embryonic form of these data gathering and measurement techniques, the firms undertaking semantic polling used a wide variety of methods to produce their data. At the simplest end of the spectrum, some firms employed simple keyword based techniques, where tweets were searched for appearances of certain words that were predefined as indicating positive or negative sentiment. This approach has obvious limitations, which will be unsurprising to any social scientists familiar with longstanding criticisms of content analysis methods (Krippendorff, 2013: 11-17). The inability to analyse context is particularly striking. As such, keyword based techniques have been abandoned by many of the firms following early experiments (social media practitioner 3).

As a result, some of our interviewees had developed proprietary software capable of undertaking natural language processing. Here, a computer essentially "reads" the content of the text and attributes a sentiment value to it (Kao & Poteet, 2007). Since the computer is not simply counting words or looking for key terms, and instead looking at the totality of the meaning of a piece of text, the context of a statement is far easier to take into account (social media practitioner 1; social media practitioner 2; social media practitioner 3). Furthermore, by allowing human beings to correct the machine's understanding of specific portions of text, the program's accuracy can be increased over time and, with it, the reliability of its analysis.

The issue of machine learning leads to another important methodological division that existed between the various semantic pollsters: the role for humans in the coding process. It should be noted that some rejected the idea outright, with one interviewee arguing that one of the virtues of this method of analysis is that it can be wholly automated and, thus, escape the possibility of human biases influencing results (social media practitioner 5). However, many interviewees acknowledged that they engaged in significant amounts of data cleaning using human coders. This reliance on human coders was said to occur because of the relatively high

failure rate of machine coding. In particular, one problem that interviewees mentioned was coding for sarcastic or ironic comments. This was seen as an intrinsic problem with computer coding methods (social media practitioner 3).

One firm in our sample was attempting to reconcile some of these ideas, experimenting with mixed data streams. This entailed a mix of three methods: Identifying highly influential Twitter users, not by number of followers but by seeing whose content and links were shared, what they termed a "depth metric"; second, identifying whether any social media user or news media outlet's 'angle' or agenda diffused; and third, correlating candidate mentions on Twitter with their poll ratings. By bringing together the analysis of conversations and voting behaviour, it was argued, "you can see a strong correlation between how parties are performing and the issues people care about" (social media practitioner 1).

These variations in semantic polling techniques are indicative of a methodology that is developing very rapidly. However, this also leads to some important questions. Media coverage of the election rarely pointed out the differences in methodologies that were being employed by different firms. Indeed, the proprietary nature of the tools meant that this information was rarely in the public domain. Key information, such as the accuracy of natural language processing techniques when compared to human coders, was never published with the results. As some scholars have pointed out, the results produced by semantic pollsters were in effect the products of black box processes, with neither journalists nor the public being given enough information to make informed judgments as to their validity (Chadwick, 2011).

Semantic polling and the traditional media

An additional explanation for the reticence of semantic pollsters when discussing their methodologies, and in particular in releasing data on the inaccuracy of their techniques, may

be found in their motivations for researching the election in the first place. A number of our interviewees admitted that they saw the election as a chance to get free media coverage for their companies and thus to raise their profile. Indeed, some went as far as to suggest that their definition of success was not making a correct prediction or a 'provable statement' about public opinion, but instead getting media coverage (social media analyst 5). Clearly, if the aim of the exercise is self-publicity of this kind, the last thing a firm will want to do is to draw attention to the weaknesses of its methodology.

This was not a one-way relationship, however. Journalists also have their own motivations for covering stories using social media as evidence for public opinion. It has been argued that media demand for social media analysis of the public has driven much of the research in the field (Panagiotis Takis Metaxas, Eni Mustafaraj, & Daniel Gayo-Avello, 2011: 2). Our interviews demonstrated that major media organisations were actually commissioning social media research to augment their election coverage (journalist 1; journalist 2). This should not surprise us, as there is a strong pre-existing tendency towards increased reporting of the opinions of the public. The most obvious evidence of this is the increased number of polls carried out in recent elections, a pattern evident in a number of countries (Brettschneider, 1997; Kellner, Twyman, & Wells, 2011; Traugott, 2005). Partially this development has been driven by public demand for increased data about the standings of candidates and parties (Iyengar, Norpoth, & Hahn, 2004). Additionally these developments may also be attributable to evolving patterns of news production. The near insatiable demand for content created by 24 hour rolling broadcast news and the development of the internet, coupled with declining journalist resources in the newsroom, have increased the centrality of polls in media coverage (Dunaway, 2011; Frankovic, 2012; Rosenstiel, 2005). In such a climate, it is hardly surprising that news organizations in the UK have taken the opportunity afforded by social media monitoring to create new measures of public opinion.

The construction of electoral narratives creates an additional connection between semantic polling and traditional media. An example of this can be seen in the phenomena of "Cleggmania", the sudden surge in both opinion poll and social media-based support for Nick Clegg in the aftermath of the first televised debate. As one of our interviewees noted, the fact that datasets from social media reinforced the impression that this development was occurring made it far more newsworthy than it otherwise would have been (pollster 4). The seeming success of Clegg among Twitter users was put down to an echo chamber effect, especially given what was argued to be the bias towards political elites and the liberal-left on Twitter in the UK (pollster 4; journalist 1). Additionally, as one of the journalists we interviewed noted, explaining the results of semantic polling is a highly subjective process (journalist 2). This makes it much more likely to be framed in the context of existing and dominant narratives.

The tendency towards dominant narratives should be no surprise, as it has been widely documented in academic studies of coverage of opinion polls, and in particular the narrow focus on the "horse race" (Larson, 2003; Patterson, 1993; Strömbäck, 2009). In the UK in 2010, it seems this pattern continued, with data from social media being integrated with traditional polling numbers to fuel the Cleggmania narrative. As one pollster we interviewed conceded, his firm's predictions began to be based more on the "Lib-Dem surge" media narrative than on their own research and that "we all got carried away with the story… we all bought in to some of that media mayhem" (pollster 2).

Differing conceptions of public opinion

Broadly, the traditional opinion pollsters in the sample offered three reactions to the possibilities offered by semantic polling: ignorance, curiosity or scepticism. Some pollsters were simply not aware of the presence of social media in political coverage, noting that it would be useful if millions of tweets could be analysed, clearly unaware that this is already

occurring (pollster 4). Some pollsters did offer positive comments on the potential of social media analysis, noting that it could offer very rapid insights into public opinion or serve specific functions, such as message testing (pollster 3; pollster 5). Perhaps unsurprisingly, however, most had a dim view of the potential of semantic polling. One very obvious explanation for this is that semantic polling might pose a threat to pollsters' livelihoods. Certainly, the nature of public opinion research is being challenged by technological developments. In 2009, a study done by the American Association of Public Opinion Research found that online research was now worth $2 billion a year. Furthermore, they estimated that 85 per cent of that revenue had been diverted from more traditional methods (AAPOR, 2009).

Beyond explanations based on self-interest, however, our interviews also demonstrated major epistemological differences between traditional pollsters and social media researchers. Traditional pollsters (including some senior figures in the self-regulatory body the British Polling Council) criticised semantic polling because it was incapable of generating a representative sample of the voting population (pollster 2; pollster 5; regulator 2; regulator 3). This was particularly reflected in a concern on the part of many of the interviewees about the demographics of the Twitter-using population as compared to the wider electorate. As one pollster argued "It doesn't help to talk to a million people if that million people are all in the same age group or the same gender, or they live in the same place or have the same political allegiance" (pollster 2).

Data contamination was an additional concern with pollsters fearing that party activists could organise online or even employ "spambots" to give a false impression of public support (pollster 2). Traditional pollsters ask a range of questions during an election campaign, examining the public's view on key issues and the popularity of would-be leaders. However,

as some interviewees noted (pollster 1; pollster 5), the ultimate measure of success for a pollster during an election campaign is predicting the outcome. This emphasis on prediction also goes a long way towards explaining criticisms of social media research as non-representative and reliant on data of dubious quality.

Advocates of social media analysis in the interview sample responded to these claims by seeking to change the terms of the debate. One interviewee overtly responded to comments about the unrepresentativeness of semantic polling by arguing that "It's not about representativeness, but it's about individuals, news organisations and their articles, and the impact they have" (social media analyst 1). Another likened their research techniques to qualitative methods, noting that it amounted to the "the biggest possible focus group" (social media analyst 6). This was said to be because the data allow for the study of interactions, not just discrete individual opinions that are aggregated.

Social media analysts argued that their techniques had certain strengths not found in other research methods. First, the data are generated without their intervention, limiting any possible distortion created by researcher intrusion (social media analyst 1; social media analyst 2). The reason this has such value, it was argued, was that there is a difference between "what people say when you ask them a direct question and what they choose to say about themselves when no one is really paying attention" (social media analyst 5). As such, semantic polling was argued to yield insights about how people talk about politics and respond to events that are not available through traditional research methods. Second, semantic polling was argued to have the potential to offer a more dynamic view of public opinion. In part, this was due to the immediacy of the data that could be gathered, seemingly offering real-time responses to political events (social media analyst 4; pollster 5). It was argued this is especially useful in the emerging hyper-mediated political environment,

wherein stories can break very rapidly and certainly more quickly than traditional opinion polls can be put in the field (pollster 5). Third, it was noted the insights into public opinion arising from semantic polling were inherently social in orientation, offering the opportunity to better understand how opinions develop through interactions and group dynamics over a prolonged period of time (pollster 5; social media analyst 5; journalist 1). Finally, on a number of occasions, interviewees noted the potential of social media to lead to greater engagement, especially on the part of young people who are least likely to partake in the political process (social media analyst 4). Here we find an interesting reversal of an argument previously made regarding the unrepresentativeness of Twitter-users by traditional pollsters. Far from being seen as a weakness, it is reported here as a corrective, since it allows a new way for previously unheard voices to enter into the political process.

<u>Discussion: public opinion theory and semantic polling</u>

Our interviews suggest a fundamental disagreement between traditional and semantic pollsters about the value of their own research methods for understanding public opinion. While those engaged in social media research highlighted distinctive virtues of their methods (especially non-interventionism, rapidity and being able to model conversations), many of the opinion pollsters we interviewed focused on the social media sample being non-representative, suggesting that semantic polling could best be characterised as an open access straw poll, a high-tech (but equally flawed) version of the Literary Digest poll (Lusinchi, 2012). The two positions are not necessarily mutually exclusive. However, in order to see if and how they can be reconciled, it is important to reconsider contemporary definitions of public opinion.

The need to develop an appropriate theoretical language to talk about semantic polling and public opinion is even more urgent now that the media is – as we document above – starting

to report social media analysis. Practitioners, reporters and citizens require the conceptual terminology to understand and explain the data being reported. A large body of research has suggested that journalists around the world lack either the inclination or the technical knowledge to explain methodological nuance. In the case of traditional polling, this is most evident in the omission of statistical information, such as margins of error (see for example: Ferguson & De Clercy, 2005; Larson, 2003; Meyer, 1968; Patterson, 2005; Pétry & Bastien, 2009; Smith & Verrall, 1985; Welch, 2002). Even when discussing qualitative public opinion methods, such as focus groups, reporters inappropriately tend to default to the language of quantitative opinion polling (Wilkinson & Kitzinger, 2000). The arguments made by traditional pollsters in our interviews make it clear that it would be equally inappropriate to talk of semantic polling in these terms. However, such an approach remains very likely in the absence of a revised definition of public opinion.

Revisiting some older concepts of public opinion

While our interviews with opinion pollsters reflect the dominance of traditional public opinion paradigm, it is not beyond the realms of possibility that the emergence of semantic polling heralds, in its very embryonic stages, the cusp of a new public opinion infrastructure. As such, we suggest that revisiting some older understandings of public opinion has the potential not only to offer a fresh perspective on how we can better understand semantic polling, but also better assess the strengths and weaknesses of this new set of methods.

James Bryce and the organs of public opinion

An instructive thinker in this regard is the nineteenth century British political theorist James Bryce. Although often cited by George Gallup as the inspiration behind his polling methods, Bryce has more recently been dismissed as outmoded and irrelevant to modern public opinion research (Bogart, 1985: 14). This is because Bryce's definition of public opinion was broader

than contemporary understandings, with the public voice being expressed through what he termed "organs of public opinion" (Bryce, 1888: chapter 92). By "organ" Bryce meant an arena where public opinion would become manifest. In his writing, Bryce identified four such organs: the press, public meetings, elections and citizen associations. Employing this approach, we might therefore think of social media as a new organ of public opinion.

Bryce's understanding of public opinion is significant for three reasons. First, the very nature of his idea of the organs of public opinion allows for multiple arenas where public opinion can become visible. In a contemporary setting, this allows opinion polling and semantic polling to co-exist as distinct manifestations of public opinion. Second, with the exception of the election itself, the environments that Bryce cites are not places where the whole electorate are represented. The technique of creating a statistically representative sample had not yet been devised at the time he was writing. However, it is also important to note that at least two of the organs of public opinion – public meetings and citizen associations – are analogous to social media, in that individuals enter into them and choose to participate if they feel inclined to do so, rather than being selected as part of a constructed sample. The assumption among the pollsters we interviewed that social media users were disproportionately likely to be politically interested, middle class and liberal-leaning reflects a similar critique. That said, there is also distinctive value in the town hall or social media-based organ of public opinion, as it allows for a more conversational definition of public opinion, something that opinion polling methods have been critiqued for neglecting (Larson, 1999). As a result, emotions, such as enthusiasm or anger, might be far more evident in a public meeting or on social media than they would appear in an opinion poll. Furthermore, while discussions online may only involve an unrepresentative and engaged subset of the population, public opinion theory would suggest that the opinions held and debates conducted by these smaller groups often pre-empt those that develop in wider society (Zaller, 1992: chapter 12).

The third point refers to the contemporary concern (evident in our interview sample, especially with the traditional pollsters) that social media might be "hijacked" by organised political parties or pressure groups. However, Bryce's idea of public opinion is inherently group-based and pluralist. The public, as it manifests itself within the organs of public opinion, reflects not only the opinions held by individuals but also the mobilisation potential of organisations, be they political parties, trade unions and pressure groups. Again we can apply this view to semantic polling. Instead of seeing such organisational capacity as a distortion – as the traditional pollsters do – a Bryce-inspired model of public opinion would instead interpret them as contours of the political landscape, and worthy of understanding. As such, if the Labour Party is able to mobilise its supporters to post #welovetheNHS, that is not a manipulation, but a political phenomenon of note.

Herbert Blumer's challenge to opinion polling orthodoxy

A second theorist of public opinion whose work could contribute to developing a conceptual framework to understand the strengths and weaknesses of semantic polling is the American sociologist Herbert Blumer.[ix] Writing at the same time as the emergence of modern opinion polling techniques, Blumer offered a very distinctive definition of public opinion which was critical of the then emerging orthodoxy. In particular, Blumer makes three claims that seem very appropriate for our assessment of the implications of semantic polling: the public is social, public opinion is hierarchical, and true public opinion requires that the public is engaged in political debates.

Blumer's first argument rejects the methodological individualism of opinion polling techniques, an orthodoxy that was expressed in the very early days of polling, based on the claim that public opinion is nothing more than the sum of individual opinions (Allport, 1937). Allport's position reflected the values of his time, and in particular assumed primacy of one-

person-one-vote electoral systems. However, Blumer rejected this position as a simplification of public opinion, neglecting its social dimension – that is, the extent to which individual opinions are generated through social interaction over time rather than in isolation (Blumer, 1948: 542-543). This argument has resonance when we consider measuring public opinion through social media, a set of techniques that can be used to evaluate the interactions between many people.

This is related to Blumer's second claim that public opinion measurement should be hierarchical. While opinion pollsters treat individuals within their sample as having equal status and their opinions as of equal import, Blumer argued that this was not how society worked. In his view, *who* held an opinion does matter, as some voices are likely to have more influence on public debate than others. Yet opinion polling gives no sense of whether a particular view was held by (in his example) "an archbishop or an itinerant laborer" (Blumer, 1951: 546). Blumer's position was that pollsters were only asking half the question. As well as knowing what people thought, analysts also needed to know the consequences of what a particular person thinks. Position in the social hierarchy and the power to influence debate become important factors. This idea clearly bears a similarity to contemporary social media analysts who are striving to measure influence online.

Blumer's third point relates to the engagement of the public and encompasses perhaps his most radical critique of opinion polling orthodoxy. Blumer argued that citizens' involvement in politics can be divided into three distinct levels depending on the type of engagement occurring: the crowd, the public and the mass (Blumer, 1951). In the first instance, citizens engage through the crowd. The crowd is a unified entity sharing the same opinion on a topic and driven by emotion (indeed, Blumer's idea of the crowd is analogous to coverage during the UK 2010 election that cited hashtags as representing a particular feeling among the

public). The mass refers to the definition of citizen engagement most closely parallel to that measured by opinion polling. People within the mass might be aware of important issues, but the nature of the opinion poll (where their opinion is solicited by a researcher in private) means that they consider them in isolation from other citizens and do not have a chance to engage in debate that would allow them to deliberate and have their opinions challenged. In short, there is no social dimension to the technique of reporting their political opinions. Blumer's definition of the public, in contrast, requires political engagement among a group of people who do not agree with each other, but are willing to debate and listen. There is much research to suggest that social media is a far from perfect environment for deliberation and high-quality political engagement (Hindman, 2008; Keen, 2007; Sunstein, 2009). However, what was occurring on Twitter during the general election arguably resembles Blumer's definition of the public more than his definition of the mass. As such, this conception offers a very different way of understanding public opinion – inherently social and conversational. Semantic polling has the potential to provide research techniques for measuring public opinion with the public being conceptualised in this way, as distinct from the definition employed by traditional pollsters.

Conclusion: the future of social media research and public opinion

Coverage of the 2010 United Kingdom election suggests that the use of social media as a tool to understand and illustrate public opinion is starting to enter into mainstream media discourse. While this occurred in a variety of ways, including electronic vox pops and commentary on trending topics, more complex semantic polling techniques are also starting to be developed and employed in election coverage. These techniques and the corporate sector that uses them are evolving very rapidly, so it seems likely they will gather more interest in the future.

What does this mean for the idea of public opinion? It is possible to measure the value of these techniques in two ways. First, their methodological effectiveness can be assessed against the existing paradigm of public opinion research, dominated by the representative sample opinion poll. Here, they are almost invariably found wanting. This might not always be the case. It is possible that future social media-based research methods will overcome this difficulty, developing weighting techniques that allow for the creation of a representative sample, in a manner similar to online panel surveys (for an exploratory attempt at employing this technique, see Sang & Bos, 2012).

While not rejecting the potential value in such innovations, this article has suggested a different approach, arguing that semantic polling, even as currently practiced, has a number of distinctive virtues. In particular, it allows public opinion to be conceptualised as being more than the sum of discrete preferences and instead as an on-going product of conversation, embedded in the social relationships that create it.

How would we conceptualise public opinion in an era of semantic polling? First, the issue of redefining public opinion goes beyond the methods available to study it. As with the representative sample opinion poll before it, semantic polling is the product of the society it exists in. The Gallupian paradigm was shaped by the ideas of the mass society of the twentieth century. Thus, the method employed aimed to discover the will of the total population, and communicate its relatively-settled preferences to the governing class. In contrast, social media research cannot claim the same level of representativeness. However, broad coverage of much of the population and ability to gather hour-by-hour or even minute-by-minute data is arguably well suited to the more restless and atomized society found in late modernity. As such, we might be able to think less about public opinion and more about the opinions of various publics, as well as how they intersect and collide.

This argument can be developed further if we consider the relationship between the different spaces in which public opinion becomes manifest and how these manifestations are understood. Traditional polling offer quantitative representations of what individuals cumulatively think, most often at the national level. In contrast, semantic polling can be used to measure and understand different organs in real time: the audience for a televised event, protesters at a demonstration, or delegates at a political conference. Furthermore, semantic polling has the potential to offer qualitative and quantitative combinations that can be read through the traditional polling prism (noting the issue of non-representativeness) or used to evaluate how influence operates in political debates and offer a range of visualisations and representations. Different organs have distinctive purposes, whether it be predicting an election winner, explaining how public opinion shifts, broadcasting striking images of public back to themselves, and so on. It is through the totality of these organs that an imaginary of the nature of public opinion is created in a society *as well as* the empirical measurement of an actual public opinion existing at that moment. Members of that public are aware of, and reflect on, both those imagined and measured publics to varying extents and interpret and talk about these forms of public opinion in their everyday lives. This assemblage evolves as new technologies and techniques emerge and as political leaders and publics develop their own conceptions of public opinion as political events develop over years and decades. This means that public opinion must thus be understood through both realist and constructivist lenses because of the reflexivity built into this infrastructure; the pressing research challenge is to understand how empirical publics are reflected on as citizens imagine public opinion *and* how those doing the measuring develop techniques framed by their own imaginaries of the nature of public opinion. This article has begun to open up these recursive loops by presenting some of the reflection going on by pollsters and journalists as they wrestle with new technologies and techniques.

Adopting a new understanding of this kind is not an unprecedented challenge: the history of public opinion research is based on a dialectic interaction between evolving theories and methods. From this perspective then, the development of semantic polling is a disruptive event, challenging us to engage in theoretical discussion and revision. The great task for public opinion researchers of all types as well the journalists who cover politics is to find ways of explaining their datasets to the audience in a way that is intelligible and nuanced, reflecting both the strengths and weaknesses of the techniques being employed. This is difficult for a number of reasons, not least because many of the social media-based methods being deployed are incredibly complex or are the product of proprietary techniques. As research on coverage of opinion polling indicates, the media does not have the greatest track record on covering this kind of detail. However, the reward is potentially very great: a richer, more complete understanding of public opinion.

Bibliography

AAPOR. (2009). *AAPOR Report on Online Panels*. Washington D.C.: American Association of Public Opinion Researchers.

Allport, F. (1937). Toward a Science of Public Opinion. *Public Opinion Quarterly, 1*(1), 7-23. doi: 10.1086/265034

Arthur, C. (2010, 3 May). Media: The election campaign: We need your Twote: This is the first British general election in the social media age - what difference has it made?, *The Guardian*.

Bauman, S., & Herbst, S. (1994). Managing perceptions of public opinion: Candidates' and journalists' reactions to the 1992 polls. *Political Communication, 11*(2), 133-144. doi: 10.1080/10584609.1994.9963020

Blumer, H. (1948). Public opinion and public opinion polling. *American Sociological Review, 13*(5), 542-549.

Blumer, H. (1951). Collective behavior. In A. M. Lee (Ed.), *New Outline of the Principles of Sociology* (pp. 67-121). New York: Barnes and Noble.

Bogart, L. (1985). *Polls and the awareness of public opinion*. New Brunswick: Transaction.

Bourdieu, P. (1979). Public Opinion Does Not Exist. In A. Mattelart & S. Siegelaub (Eds.). *News and the Empowerment of Citizens* (pp.124-130). New York: International General,

Boyatzis, R. E. (1998). *Transforming qualitative information: Thematic analysis and code development*. London; Thousand Oaks, CA: Sage.

Brettschneider, F. (1997). The Press and the Polls in Germany, 1980-1994: Poll Coverage as an Essential Part of Election Campaign Reporting. *International Journal of Public Opinion Research, 9*(3), 248-265. doi: 10.1093/ijpor/9.3.248

Bryce, J. B. (1888). *The American Commonwealth*. London: Macmillan.

Cellan-Jones, R. (2010a, 22 April). 'It's all Nick Clegg's fault'. *dot.Rory.* Retrieved 25 April, 2012, from

http://www.bbc.co.uk/blogs/thereporters/rorycellanjones/2010/04/its_all_nick_cleggs_fault.html

Cellan-Jones, R. (2010b, 30 April). Online 'sentiment' around the prime-ministerial debates. *dot.Rory.* Retrieved 23 April, 2012, from

http://www.bbc.co.uk/blogs/thereporters/rorycellanjones/2010/04/online_sentiment_around_the_pr.html

Chadwick, A. (2011). Britain's First Live Televised Party Leaders' Debate: From the News Cycle to the Political Information Cycle. *Parliamentary Affairs, 64*(1), 24-44. doi: 10.1093/pa/gsq045

Childs, H. (1939). By Public Opinion I Mean. *Public Opinion Quarterly, 3*(2), 327-336. doi: 10.1086/265298

Dunaway, J. (2011). Poll Centred News Coverage: Causes and Consequences. In K. Goidel (Ed.), *Political Polling in the Digital Age : The Challenge of Measuring and Understanding Public Opinion* (pp. 71-84). Batin Rouge: Louisana State University Press.

Elmer, G. (2013). Live research: Twittering an election debate. *New Media & Society, 15*(1), 18-30. doi: 10.1177/1461444812457328

Ferguson, P. A., & De Clercy, C. (2005). Regulatory compliance in opinion poll reporting during the 2004 Canadian election. *Canadian Public Policy/Analyse de politiques, 31*, 243-257.

Franch, F. (2013). (Wisdom of the Crowds) 2: 2010 UK election prediction with social media. *Journal of Information Technology & Politics, 10*(1), 57-71. doi: 10.1080/19331681.2012.705080

Frankovic, K. A. (2012). Opinion Polls and the Media in the United States. In C. Holtz-Bacha & J. Strömbäck (Eds.), *Opinion Polls and the Media: Reflecting and Shaping Public Opinion* (pp. 113-133). Basingstoke, UK: Palgrave Macmillan.

Gallup, G. H. (1939). Public opinion in a democracy. Princeton, NJ: Herbert L. Baker foundation, Princeton University Press.

Gayo-Avello, D. (2011). Don't turn social media into another 'Literary Digest'poll. *Communications of the ACM, 54*(10), 121-128. doi: 10.1145/2001269.2001297

Gayo-Avello, D. (2013). A meta-analysis of state-of-the-art electoral prediction from Twitter data. *Social Science Computer Review, 31*(6), 649-679. doi: 10.1177/0894439313493979

Glaser, B. G., & Strauss, A. L. (1968). *The discovery of grounded theory : strategies for qualitative research*. London: Weidenfeld and Nicolson.

Gonzalez-Bailon, S., Banchs, R. E., & Kaltenbrunner, A. (2010). Emotional reactions and the pulse of public opinion: Measuring the impact of political events on the sentiment of online discussions. *arXiv preprint arXiv:1009.4019*.

Greenslade, R. (2010, 26 April). Media: Opinion: Politics and the media: Will it be the Sun or Twitter wot won it?, *The Guardian*.

Guest, G., MacQueen, K. M., & Namey, E. E. (2011). *Applied Thematic Analysis*. London; Thousand Oaks, CA: Sage.

Herbst, S. (1993). *Numbered voices : how opinion polling has shaped American politics*. Chicago: University of Chicago Press.

Herbst, S. (2001). Public Opinion Infrastructures: Meanings, Measures, Media. *Political Communication, 18*(4), 451-464. doi: 10.1080/10584600152647146

Herbst, S. (2011). (Un)Numbered Voices. In K. Goidel (Ed.), *Political Polling in the Digital Age: The Challenge of Measuring and Understanding Public Opinion* (pp. 85-98). Batin Rouge: Louisana State University Press.

Hindman, M. (2008). *The Myth of Digital Democracy*. Princeton, N.J.: Princeton University Press.

Hinsliff, G. (2010, 11 April). Web 2.0: the new election superweapon, *The Observer*.

Iyengar, S., Norpoth, H., & Hahn, K. (2004). Consumer Demand for Election News: The Horserace Sells. *Journal of Politics, 66*(1), 157-175. doi: 10.1046/j.1468-2508.2004.00146.x

Jensen, M., & Anstead, N. (2013). Psephological Investigations: Tweets, Votes, and Unknown Unknowns in the Republican Nomination Process. *Policy and Internet, 5*(2), 23-44. doi: 10.1002/1944-2866.POI329

Jungherr, A., Jürgens, P., & Schoen, H. (2012). Why the Pirate Party Won the German Election of 2009 or The Trouble With Predictions: A Response to Tumasjan, A., Sprenger, T. O., Sander, P. G., & Welpe, I. M. "Predicting Elections With Twitter: What 140 Characters Reveal About Political Sentiment". *Social Science Computer Review, 30*(2), 229-234. doi: 10.1177/0894439311404119

Kao, A., & Poteet, S. R. (2007). *Natural Language Processing and Text Mining*. London: Springer.

Keen, A. (2007). *The Cult of the Amateur*. New York: Currency.

Kellner, P., Twyman, J., & Wells, A. (2011). Polling Voting Intentions. In D. Wring, R. Mortimore & S. Atkinson (Eds.), *Political communication in Britain TV debates, the media, and the election* (pp. 94-108). Houndmills, Basingstoke, Hampshire ; New York, NY: Palgrave Macmillan.

Krippendorff, K. (2013). *Content analysis : an introduction to its methodology* (3rd ed.). Los Angeles ; London: SAGE.

Larson, S. G. (1999). Public opinion in television election news: Beyond polls. *Political Communication, 16*(2), 133-145. doi: 10.1080/105846099198695

Larson, S. G. (2003). Misunderstanding Margin of Error Network News Coverage of Polls during the 2000 General Election. *The International Journal of Press/Politics, 8*(1), 66-80. doi: 10.1177/1081180X02238785

Lewis, J. (2001). *Constructing Public Opinion: How Political Elites do what they like and why we seem to go Along with it*. New York: Columbia University Press.

Lewis, J., Wahl-Jorgensen, K., & Inthorn, S. (2004). Images of citizenship on television news: constructing a passive public. *Journalism Studies, 5*(2), 153-164. doi: 10.1080/1461670042000211140

Lusinchi, D. (2012). "President" Landon and the 1936 Literary Digest Poll: Were Automobile and Telephone Owners to Blame? *Social Science History, 36*(1), 23-54. doi: 10.1215/01455532-1461650

Metaxas, P. T., Mustafaraj, E., & Gayo-Avello, D. (2011). *How (Not) to Predict Elections*. Paper presented at the Privacy, security, risk and trust (passat), 2011 ieee third international conference on and 2011 ieee third international conference on social computing (socialcom), Boston, MA.

Metaxas, P. T., Mustafaraj, E., & Gayo-Avello, D. (2011). *Use this one: How (Not) to Predict Elections*. Paper presented at the Privacy, security, risk and trust (passat), 2011 ieee third international conference on on social computing (socialcom), Boston, MA.

Meyer, P. (1968). Truth in Polling. *Columbia Journalism Review, 7*(2), 20-23.

Moon, N. (1999). *Opinion polls: history, theory and practice*. Manchester: University of Manchester Press.

O'Connor, B., Balasubramanyan, R., Routledge, B. R., & Smith, N. A. (2010). *From Tweets to Polls: Linking Text Sentiment to Public Opinion Time Series*. Paper presented at the ICWSM, Washington, DC.

Parry, K., & Richardson, K. (2011). Political Imagery in the British General Election of 2010: The Curious Case of 'Nick Clegg'. *The British Journal of Politics & International Relations, 13*(4), 474-489. doi: 10.1111/j.1467-856X.2011.00452.x

Patterson, T. E. (1993). *Out of order*: Knopf New York.

Patterson, T. E. (2005). Of polls, mountains US journalists and their use of election surveys. *Public Opinion Quarterly, 69*(5), 716-724. doi: 10.1093/poq/nfi065

Pétry, F., & Bastien, F. (2009). *From Pollsters to Journalists: Inaccuracies in Horse-race Coverage During the 2008 Canadian Election*: Center for the Analysis of Public Policy.

Roberts, B. (2010, 23 April). It's all Nick Clegg's fault: Wars... Famine... That poor bloke who needs a kidney transplant, *The Daily Mirror*.

Robinson, N. (2012). *Live from Downing Street*. London: Bantam.

Rosenstiel, T. (2005). Political Polling and the New Media Culture: A Case of More Being Less. *Public Opinion Quarterly, 69*(5), 698-715. doi: 10.1093/poq/nfi062

Rowlett, J. (Presenter). (2010, 29 April). BBC Newsnight [Television]. London: BBC.

Sang, E. T. K., & Bos, J. (2012). *Predicting the 2011 dutch senate election results with twitter.* Paper presented at the Proceedings of the Workshop on Semantic Analysis in Social Media, Stroudsburg, PA.

Sawyer, M. (2010, 17 April). MUCH hoo-ha about this being the Twitter election, *Daily Mirror*.

Shamma, D. A., Kennedy, L., & Churchill, E. F. (2009). *Tweet the debates: understanding community annotation of uncollected sources.* Paper presented at the Proceedings of the first SIGMM workshop on Social media.

Smith, T. J., & Verrall, D. O. (1985). A critical analysis of Australian television coverage of election opinion polls. *Public Opinion Quarterly, 49*(1), 58-79. doi: 10.1086/268901

Splichal, S. (2012). Public Opinion and Opinion Polling: Contradictions and Controversies. In C. Holtz-Bacha & J. Strömbäck (Eds.), *Opinion Polls and the Media: Reflecting and Shaping Public Opinion* (pp. 25-48). Basingtstoke: Palgrave Macmillan.

Stromback, J. (2012). The Media and Their Use of Opinion Polls: Reflecting and Shaping Public Opinion. In C. Holtz-Bacha & J. Strömbäck (Eds.), *Opinion Polls and the Media: Reflecting and Shaping Public Opinion*. Basingtstoke: Palgrave Macmillan.

Strömbäck, J. (2009). Vox Populi or Vox Media?: Opinion Polls and the Swedish Media, 1998-2006. *Javnost-The Public, 16*(3), 55-70.

Sudhaman, A. (2010). Lib Dems' online sentiment surges, PRWeek's General Election sentiment tracker shows. *PR Week*.

Sullivan, A. (2010, 25 April). Clegg catches the angry transatlantic wind, *The Sunday Times*.

Sunstein, C. R. (2009). *Republic.com 2.0*. Princeton, NJ: Princeton University Press.

Thorpe, V. (2010, 18 April). Television satirists struggle to identify targets in the kingdom of the bland, *The Observer*.

Traugott, M. W. (2005). The Accuracy of the National Preelection Polls in the 2004 Presidential Election. *Public Opinion Quarterly, 69*(5), 642-654. doi: 10.1093/poq/nfi061

Tumasjan, A., Sprenger, T. O., Sandner, P. G., & Welpe, I. M. (2010). Predicting Elections with Twitter: What 140 Characters Reveal about Political Sentiment. *Word. Journal Of The International Linguistic Association*, 178-185.

Tumasjan, A., Sprenger, T. O., Sandner, P. G., & Welpe, I. M. (2011). Election forecasts with Twitter how 140 characters reflect the political landscape. *Social Science Computer Review, 29*(4), 402-418. doi: 10.1177/0894439310386557

Twitterfall. (2012). Twitterfall.com.   Retrieved 10 July 2012, from http://twitterfall.com/

Urquhart, C. (2013). *Grounded Theory for Qualitative Research: a Practical Guide*. London; Thousand Oaks, CA: Sage.

Wallsten, K. (2011). *Micrblogging and the news: Political elites and the ultimate retweet*. Paper presented at the American Political Science Association, Seattle, WA.

Ward, V. (2010, 24 April). Daddy, You've Won! Son's Message to Lib Dem Star, *Daily Mirror*, p. 2.

Welch, R. L. (2002). Polls, Polls, and More Polls An Evaluation of How Public Opinion Polls Are Reported in Newspapers. *The Harvard International Journal of Press/Politics, 7*(1), 102-114. doi: 10.1177/1081180X0200700107

Wilkinson, S., & Kitzinger, C. (2000). "Clinton faces nation": a case study in the construction of focus group data as public opinion. *The Sociological Review, 48*(3), 408-424. doi: 10.1111/1467-954X.00223

Zaller, J. (1992). *The Nature and Origins of Mass Opinion*. Cambridge: Cambridge University Press.

Table One: Uses of social media in 2010 election coverage and conceptions of the public

Social media data use	Historic parallel	Conception of the public
Quotation of individual status update	Vox pop interview	Individual as representative of strand in public opinion, "person on the street"
Raw quantitative data (quoting number of social media updates during events; trending topics on social media)	Accounts of election turnout; mobilisation in favour of cause, position	Unified reaction among public or subset of public, linked to specific emotional disposition (apathy, engagement, anger etc)
Semantic polling	Straw poll; opinion poll	Seeks to quantify a divided public

Appendix One: Anonymous list of interviewees

No.	Citation ID	Professional group	Name of company	Date of interview
1	Socialmedia 1	Socialmedia analytics	UK-based, politically focused socialmedia analytics firm	04-Jan-11
2	Socialmedia 2	Socialmedia analytics	Major UK-based socialmedia analytics firm	04-Aug-11
3	Pollster 1	Pollster	Major UK-based online pollster	05-Aug-11
4	Socialmedia 3	Socialmedia analytics	International socialmedia analytics firm	05-Aug-11
5	Party 1	Political party	Major UK political party	10-Aug-11
6	Regulator 1	Regulator	Electoral Commission	11-Aug-11
7	Pollster 2	Pollster	International online pollster	29-Aug-11
8	Socialmedia 4	Socialmedia analytics	UK-based, marketing focused socialmedia analytics firm	01-Sep-11
9	Journalist 1	Journalist	BBC	06-Sep-11
10	Socialmedia 5	Socialmedia analytics	UK-based, marketing focused socialmedia analytics firm	12-Sep-11
11	Party 2	Political party	Major UK political party	14-Sep-11
12	Party 3	Political party	Major UK political party	14-Sep-11
13	Party 4	Political party	Major UK political party	14-Sep-11
14	Pollster 3	Pollster	International online pollster	09-Nov-11
15	Socialmedia 6	Socialmedia analytics	Major UK-based socialmedia analytics firm	19-Nov-11
16	Regulator 2	Regulator / pollster	British Polling Council	01-Dec-11
17	Pollster 4	Pollster	Major UK-based pollster	01-Dec-11
18	Pollster 5	Pollster	Major UK-based pollster	12-Dec-11
19	Journalist 2	Journalist	Major UK-based news magazine	01-Feb-12
20	Regulator 3	Regulator	British Polling Council	25-Feb-12

Endnotes

---

[i] [Anonymised project] also gave [Author 2] the chance to study other case studies such as the 2010 Haiti earthquake and 2009 UK swine flu vaccine campaign, and was designed to understand how citizens use social media to communicate in unfolding events. [Author 2] gained first-hand experience of designing, constructing and conducting social media analysis and disseminating the results to national news media immediately after events.

[ii] These data was gathered by searching for articles that referenced both "Twitter" and "election" and were published during the statutory election campaign period (6th April 2010 – 6th May 2010). Nexis allowed us to search the following British national newspapers (number of articles gathered in parenthesis): The Guardian (49); The Daily Star (7); The Morning Star (2); The Daily Mail (15); The Independent (22); The Mirror (10); The Sun (11); The Times (20); The Daily Telegraph (37); The Express (10); Independent on Sunday (4); The Mail on Sunday (2); The Observer (15); The Sunday Times (12); The Sunday Telegraph (6); Sunday Mirror (1); The Sunday Express (4).

[iii] We would like to thank our two research assistants, [Anonymous RA1] and [Anonymous RA2] for conducting and transcribing the interviews, and also the [Anonymous University 1] for providing the financial support for this research.

[iv] In the interests of transparency, we would be happy to share anonymised interview transcripts with interested researchers. Please contact the authors directly for access.

[v] Broken down, this subset of the data comprises of: 3 sections on the BBC Newsnight programme; 9 blog entries published on television channel websites (BBC and Channel 4); 25 print media stories; and 10 additional sources including press releases from social analysis firms and articles in specialist publications such as Adweek.

[vi] Rowlett actually misspoke here. The numbers cited are based on data gathered by the political social media monitoring firm Tweetminster, who actually claimed 154,000 debate related tweets published during the third debate.

---

vii See also Cellan-Jones 2010a; Greenslade 2010; Sullivan 2010; Ward 2010. For post-election academic analysis of this event, see Chadwick 2011a; Parry and Richardson 2011.

viii Although, it should be noted that some solicitation of opinions from the public does occur. Many television programmes now encourage viewers to share their views about the broadcast by publicising a hashtag, for example.

ix It should be noted that we are not the first to attempt to use Blumer's work to understand social media and public opinion. Susan Herbst (2011) argued he would have likely approved of blogs as a mechanism for better understanding the public.