Trending Twitter Topics in English: An International Comparison[1]

David Wilkinson, Mike Thelwall

Statistical Cybermetrics Research Group, School of Technology, University of Wolverhampton, Wulfruna Street, Wolverhampton WV1 1LY, UK.

E-mail: d.wilkinson@wlv.ac.uk, m.thelwall@wlv.ac.uk

Tel: +44 1902 321452, +44 1902 321470, Fax: +44 1902 321478

The worldwide span of the microblogging service Twitter gives an opportunity to make international comparisons of trending topics of interest, such as news stories. Previous international comparisons of news interests have tended to use surveys and may bypass topics not well covered in the mainstream media. This article uses 9 months of English tweets from the UK, USA, India, South Africa, New Zealand and Australia. Based upon the top 50 trending keywords in each country from the 0.5 billion tweets collected, festivals or religious events are the most common, followed by media events, politics, human interest and sport. US trending topics have the most interest in the other countries and Indian trending topics the least. Conversely, India is the most interested in other countries' trending topics and the US the least. This gives evidence of an international hierarchy of perceived importance or relevance with some issues, such as the international interest in US Thanksgiving celebrations, apparently not being directly driven by the media. This hierarchy echoes, and may be caused by, similar news coverage trends. Whilst the current imbalanced international news coverage does not seem to be out of step with public news interests, the political implication is that the Twitter-using public reflects, and hence seems to implicitly accept, international imbalances in news media agenda setting rather than combatting them. This is a problem for those believing that these imbalances make the media too powerful.

Introduction

International differences in news coverage and peoples' news interests are important not only for the media deciding what content to produce and whether to customise it for different geographic audiences but also because of the influence of the media in deciding upon the main issues of the day ("agenda setting", as discussed below). Current research into international differences in news interests typically uses surveys, but the social web has created possibilities to address this issue by analysing social media texts instead. This different angle may give new insights, such as topics of interest that are not well covered in the mainstream media. From the social web, Twitter is a particularly relevant source for news-related information because of its emphasis on rapidly sending short messages.

At the time of writing, Twitter was the ninth most visited web site in the world according to Alexa (2011) even though it can be used via mobile phones, bypassing Alexa's data collection methods. Twitter allows users to publically broadcast short text messages of up to 140 characters, known as tweets. The widespread international adoption of Twitter has created a situation in which the comments of a large number of people, albeit a small minority of the world's population, are publically available for research to see what kinds of things generate interest and even how topics of interest vary by country. These are important because Twitter is a phenomenon in its own right and because of the potential to gain insights into underlying human interests and international differences in reactions to the news and political issues. Twitter itself promotes this kind of analysis by publishing keyword trends on users' profile pages and even annual reviews of the top tweeted topics (Twitter, 2011), as do some other social media sites to a limited extent, such as Facebook (Hernandez, 2011).

The public availability of Twitter data has allowed much research into its uses (e.g., Java, Song, Finin, & Tseng, 2007; Naaman, Becker, & Gravano, 2011) and exploitation for purposes such as health research (Heaivilin, Gerbert, Page, & Gibbs, 2011; Signorini, Segre, & Polgreen, 2011), marketing (Jansen, Zhang, Sobel, & Chowdury, 2009), government information (Wigand, 2010),

---

political predictions (Tumasjan, Sprenger, Sandner, & Welpe, 2010), identifying global mood changes (Golder & Macy, 2011), technology development (Pak & Paroubek, 2010), detecting or recommending news stories (Abel, Gao, Houben, & Tao, 2011; Jackoway, Samet, & Sankaranarayanan, 2011; Phelan, McCarthy, & Smyth, 2009), identifying crowded places (Fujisaka, Lee, & Sumiya, 2010), predicting future citations (Eysenbach, 2011) or even identifying earthquake sites (Sakaki, Okazaki, & Matsuo, 2010). Nevertheless, although some articles have studied aspects of trending topics (Naaman et al., 2011; Thelwall, Buckley, & Paltoglou, 2011), none have conducted an international comparison or identified the types of topics that attract national interest. This is an important omission because international flows of information are of political and media interest. Much research, including some reviewed below, indicates that there are international inequalities in the provision of news, partly due to practical issues like the availability of convenient sources. There is less research into the public reception of such news: do inequalities in coverage lead to inequalities in interests or a reaction against the inequalities? And are these inequalities reflected or combatted in Twitter, which partly plays the role of an informal news source?

This article addresses the above questions for a special case: Tweets in English that can be geolocated to identify their region of origin. Since the typical tweet may be personal in orientation, the focus is on the most commonly tweeted news-related issues rather than more mundane communications. Hence the analysis is based upon the top trending keywords in each country with the goal of investigating international differences in news interests. The restriction to English is a practical one to minimise the impact of language differences on the results.

Literature Review

This review covers two separate topics that are needed to underpin the research. First, research into international news coverage and reception forms the main theoretical basis for comparisons with Twitter. Second, media agenda setting theory provides a basis from which to evaluate the wider significance of the findings. Finally, background information about Twitter gives overall context to the data used.

International differences in news provision and reception

Previous research suggests that trending topics in Twitter tend to be news stories (Thelwall, Buckley, & Paltoglou, 2011). Since this article focuses on international differences in trending topics in Twitter, it is useful to understand something of the role of the nation in news reporting. First, it seems to be common sense that individuals will tend to be more interested in news that is geographically closer to them: hence the proliferation of regional and national newspapers rather than international media. Nevertheless, there are various news organisations that attempt to provide international services, such as the BBC, CNN, Deutche Welle, and Al Jazeera, as well as international news agencies, such as Reuters and RIA Novosti and international news portals like PressEurop. The internet also provides access to international news sources and other countries' national news (Ahlers, 2006), which seems to give the young enhanced access to foreign news and foreign news sources (Kwak, Poor, & Skoric, 2006). Access to foreign news sources may have particular relevance to migrant populations (Kissau & Hunger, 2008) and so international mobility may also spread news stories and help foreign news and news frames to filter into a given country. Nevertheless, accessing other countries' news sources seems to happen most outside the richer nations (Berger, 2009) which presumably means that richer nations' news is more known inside poorer nations than vice versa.

There are some systematic reasons why news from one country is covered in another. These include the extent to which the two nations interact economically, cultural similarities, and geographic proximity (Kolmer & Semetko, 2010; Wu, 2000). In addition, larger countries' news seems to be more likely to be covered elsewhere and the US is particularly well discussed everywhere (Lin, Lo, & Wang, 2011; Wu, 2000), presumably because of its cultural industry and its worldwide political influence. Conversely, the US seems to favour international news from geographically close countries, as well as news that is unusual, relevant to the US, or has the potential for social change (Chang, Shoemaker, & Brendlinger, 1987). More generally, foreign news in the US has been dominated by stories about economically advanced nations (Chang, 1998), although countries in which the US is involved in conflicts, such as Iraq and Afghanistan (and Vietnam [Larson, 1982] in

the past), are presumably exceptions. For example, "despite presence of wide-scale famine, civil conflict, disputed elections and an AIDS epidemic, the African continent received limited coverage" during 2002-4 in the US, with trade ties and Gross Domestic Product being the two key predictors of coverage for African nations (Golan, 2008, p. 41; see also Besova, & Cooley, 2009).

Media agenda setting and news framing

Research in media studies and journalism has introduced the theory of *agenda setting* (McCombs, & Shaw, 1972) and for this theory the various different media (e.g., TV, print, online) are probably interchangeable (e.g., Golan, 2006). Agenda setting theory essentially postulates that the media is highly successful at telling the public *which* issues of the day are relevant. As discussed below, there is much evidence of a correlation between public awareness and media coverage, and some research has also demonstrated experimentally that media coverage can *cause* public interest in a topic (e.g., Iyengar, Peters, & Kinder, 1982). This suggests that newsroom decisions about which stories, or which aspects of stories, to publish, promote or investigate transfer to the public consciousness into a belief about what is relevant (McCombs & Bell, 1996). This has implications for the international interests of the public due to the inequalities in international news coverage discussed above. For instance, a US survey found a strong correlation between the countries that people thought were "vital to US interests" and the extent of media coverage of those countries – although some oil rich countries received little media coverage at the time (e.g., Kuwait) but were nevertheless perceived to be important, and India was perceived as relatively unimportant for the amount of coverage it received (Wanta, Golan, & Lee, 2004).

Agenda setting can occur internationally as well as nationally, due to international sources like the BBC World Service and CNN International. The latter, for example, although it customises its international news for different world regions rather than delivering US content worldwide, covers similar topics everywhere (Groshek, 2008). Moreover, a study of a range of different world news sources found all of their coverage to be similar (Loomis, 2009), suggesting that there is a global news agenda, although it is not clear which factors, if any, drive it. Another way in which framing occurs at the international level is that stories about other countries can be given a slant for the reporting nation, such as by explicitly mentioning the story's relationship to it (Wanta, Golan, & Lee, 2004). International news can also be expected to be given a national "framing", focus or orientation (Entman, 1993; Matthes, 2009), perhaps differing in each country reported (Riegert, 2011), and influencing public perceptions of the event covered (e.g., Huang, 2010).

Agenda setting theory seems to be widely accepted at the moment (e.g., Groshek, 2008), but more controversial is the stronger belief that the media is also *sometimes* influential at persuading the public what *attitude* to take about relevant issues (Wanta, Golan, & Lee, 2004). There is evidence, for example, that negative coverage of foreign nations in US news associates with negative public perceptions of those countries (Besova, & Cooley, 2009; Wanta, Golan, & Lee, 2004).

From a policy perspective, international news selection can also have important political implications because news media may simply ignore stories that do not fit comfortably within their national self-perception or that show the host nation in a poor light (Herman & Chomsky, 1988). Media agenda setting has political implications because of its relationship with foreign policy (Bennett, Lawrence, & Livingston, 2007; Cohen, 1963; Livingston, 1992; Zaller & Chiu, 1996), making this issue particularly important. For example, US news reporting of two similar airline shootings different substantially and in line with its geopolitical goals: one described as a moral outrage and the other as a technical problem (Entman, 1991). Another example is the low 1990s coverage of East Timor in the US news, where it seemed to be difficult to present US foreign policy in a good light (Chomsky, 2000). In fact, some claim that the free press is frequently used, sometimes in subtle ways, by those with influence (and not always successfully, as in the case of Hurricane Katrina coverage: Entman, 1994) to "manufacture consent" in the population to align them to the goals of the powerful (Herman & Chomsky, 1988). In particular, the type and extent of media coverage can influence the success or failure of specific foreign policy initiatives, such as wars (e.g., see the claims in: Chomsky, 2000; Der Doran, 2009; Kalb, 1994). Whilst most research into these issues has a national focus, the existence of powerful international news organisations and the

international factors in news awareness discussed above give it an international dimension (e.g., Groshek, 2008).

Topics and trends in Twitter

The ways in which people use Twitter may relate to the kind of content posted and hence the issues that emerge from an analysis of trending topics. An early small qualitative study suggested that Twitter was used for informal social interactions (D. Zhao & Rosson, 2009). This was corroborated by a later study of 317 users which found that people needing to informally connect with others were more frequent Twitter users (Chen, 2011). The idea that Twitter is used for interaction rather than just broadcasting was supported by research which found that messages targeted at individuals were often (31% of the time) responded to (Honeycutt, & Herring, 2009). Moreover, although tags seem to be used in most systems to organise content or to aid information retrieval (e.g., Dotan & Zaphiris, 2010), Twitter hashtags tend to help organise conversations rather than content (Huang, Thornton, & Efthimiadis, 2010). Hence there is a wide range of types of information that support the thesis that Twitter is a space supporting significant social interactions. Despite this, however, people may also use it for specific information needs (e.g., Hughes & Palen, 2009) or for commercial reasons (Jansen, Sobel, & Cook, 2011; Zhang, Jansen, & Chowdhury, 2011), and so it is not a purely social space.

A wide variety of types of tweet are posted, only some of which seem likely to be present in trends. In April, 2009, about 80% of users tended to post about themselves, promoting or describing their current activities, whereas 20% produced more informational postings (Naaman, Boase, & Lai, 2010, see also Dann, 2010). This study characterised tweets with nine types of topic: information sharing, self-promotion, opinions/complaints, statements and random thoughts, me now, questions, presence maintenance (e.g., "I'm back"), anecdote (me), and anecdote (others). Similarly, an analysis of 7040 tweets from as many users in November 2009 (Westman & Freund, 2010) produced five genres: personal updates, directed dialogue, real-time sharing (news and information), business broadcasting, information seeking. One study has assessed the extent to which Twitter users post about the news, using a comparison with the New York Times. The largest difference found was that the topic category *Family and Life*, presumably including the *me now* and *personal updates* categories above, accounted for about 27% of extracted Twitter topics but no topics from the New York Times (W. X. Zhao et al., 2011).

In terms of trending topics, it seems that the information sharing category in the first study and the real-time sharing category in the second study seem to be the most likely to be news related and hence to generate trends because news stories are broadcast within a relatively narrow time frame. In fact, it seems that the majority of trends in Twitter are news-related (Kwak, Lee, Park, & Moon, 2010). Moreover, Twitter tends to be used to comment on news stories rather than creating them (or merely announcing them) (Subašić & Berendt, 2011) and so it can be expected to co-ordinate well with mainstream media news (for Twitter uses by journalists see Farhi, 2009). It seems likely, however, that the mainstream media would lead Twitter except perhaps in cases where there is mass involvement in a news event, such as a disaster, and the media use Twitter as a significant source of information (e.g., Farhi, 2009; Hermida, 2010). Nevertheless, the *me now/personal updates* category may also co-ordinate and produce a trend when many people are doing the same thing, such as attending a mass rally or concert, or doing something specific to a holiday or festival day. Hence trends in Twitter may not relate solely to specific news stories. An important topic not categorised in the above research, however, is Spam (Wang, 2010), which can presumably cause trends, if not filtered out.

As introduced above, the typical topics in Twitter posts seem unlikely to be the same as the common types of trending topics in Twitter. Topics and trends in Twitter have been previously analysed for a single geographic location: New York between September 2009 and March 2010, as self-identified by Twitter users on their profiles (Naaman et al., 2011). In this study, trends published by Twitter for New York and trends identified with a time series method from the raw data were compared using quantitative and qualitative approaches. For the qualitative approach, 50 trends were categorised into four exogenous types (broadcast media events, global news events, national holidays and memorial days, and local participatory and physical events) and three endogenous types (memes,

retweets, and fan community activities). Comparing these categories to those discussed above for all tweets confirms that typical trend topics are different from typical overall topics in Twitter.

Some research has focused on trends in Twitter to understand how tweets within a trend change over time. A study of the role of sentiment in the top external trending topics in a month of Twitter found that spiking topics were typically associated with small increases in negative sentiment (Thelwall et al., 2011). A more fine-grained study of Tweets related to US college campus shootings (i.e., single broad news stories), showed that the role of Twitter changed over time, from information sharing at the start to opinion sharing later on (Heverin & Zach, 2012).

An important aspect of trending in Twitter is that information can flow through follower networks (Lerman, & Ghosh, 2010) by popular tweets being retweeted, by users responding to others' tweets (Bruns, in press; Romero, Meeder, & Kleinberg, 2011), or by repeated exposure to new hashtags (Romero et al., 2011). Since Twitter is international, this gives the potential for information to flow across international borders. This phenomenon was investigated in a study of 1,953 users sampled from 0.5 million Tweets gathered over a week in August 2009, together with a randomly selected Twitter user that they Followed (Takhteyev, Gruzd, & Wellman, 2012). Tweeter locations were identified primarily by coding the context of tweets sent, rather than by using embedded geocodes, giving greater coverage than the current paper (see methods section below). National locations were determined for all these users and regional locations for 1,259 pairs. The analysis was based upon comparing the locations of the Follower/Followee pairs, and showed that people tend to follow others if they are geographically close, and that country and language are important in predicting ties. Hence national borders will restrict the flow of information in Twitter to some extent because international connections are relatively less frequent than national ones.

In summary, existing research has shown that despite the typically quite personal content of tweets, trends in Twitter tend to be related to the news. It would therefore be possible to investigate international differences in news interests through large-scale investigations of Twitter. From the previous sub-section, it is already known that there are systematic international differences in news interests but it is not known whether these also occur in Twitter and, if so, whether these differences echo offline differences or are modified by the medium in some way.

Research questions

As shown in the literature review, trending topics in Twitter are likely to be predominantly related to news events and an analysis of them may therefore illuminate the news interests of the public. This study attempts to identify international differences in trending topics in Twitter to assess the extent of these international differences and to see whether these align with what is known from media research, and agenda setting theory in particular. Whilst this theory has been tested before, this has typically been through questionnaires or opinion polls and Twitter provides a new angle on the issue. For practical reasons, discussed below, the data is limited to Twitter users, which skews the findings towards richer and perhaps more technological users, especially in poorer nations. The goal is therefore restricted to seeking differences in this biased sample. This is exploratory research using the information-centred research approach (Thelwall, Wouters, & Fry, 2008). The objectives are therefore data driven rather than theory driven: to discover what kinds of information (in this case trending topics) can be found in Twitter, and what kind of insights the results can give on the international differences in news coverage discussed above. In particular, the main goal is to assess how trending topics (and hence Twitter-mediated news interests) vary by country rather than to test specific hypotheses. This gives the following research questions.

- Q1: What are the main *types* of trending topics for each country?
- Q2: How do the specific trending topics vary by country?

Based upon the discussions of news reception in the literature review, however, it seems reasonable to make some hypotheses relating to the second research question about the extent of international connections in the data.

- **H2a**: Trending topics of interest to the USA will be of most interest to other countries and trending topics of interest to India will be of least interest to other countries. This is because trending topics of interest within any particular country are likely to be dominated by topics

related to that country and the research discussed above suggests that economically advanced countries' news is of more external interest than poorer countries' news.

- **H2b**: The USA (as a world leading nation) and India (because of cultural differences) will be least interested in trending topics of interest from other countries.

The concept of *trending topic* is operationalized heuristically and with reference to external events rather than purely quantitatively in this article. A previous article used the term *trend* instead, "a trend on Twitter (sometimes referred to as a trending topic) consists of one or more terms and a time period, such that the volume of messages posted for the terms in the time period exceeds some expected level of activity (e.g., in relation to another time period or to other terms)" (Naaman et al., 2011). This definition is adopted except that the "one or more terms" must be related to an external event and this external event must be broadly the same for each of the terms associated with a given trending topic. The tie to an external event excludes long running topics and interests, such as, chess or Justin Bieber, except when they are tied to specific events, such as a competition or a concert. In these cases the trending topic would be the event itself (i.e., the competition or concert) rather than the long running topic or interest.

Methods

The research design was to collect and analyse texts from Twitter over an extended period in a single language across different locations to allow an international comparison. Although inter-language connections can be important on the web (e.g., Hale, 2012) the use of a single language minimises the impact of linguistic issues on the results (e.g., Nakasaki, Kawaba, Utsuro, & Fukuhara, 2009), which would otherwise be a significant factor for the automatic trending topic identification stage. English was chosen because this language is used in enough different countries to give a useful sample. Other languages used to some extent in many countries include Spanish, French and Chinese but English has the advantage of being the most common second language (Crystal, 1997). Nevertheless, the use of a single language is a limitation and different results may have been obtained for other languages.

As described below, a time series analysis method was used to extract the top 50 trending topics for each country and then a content analysis was used to identify the types of topics represented. The types of topics represented as well as the individual topics were then compared across nations.

Data

Twitter was sampled from seven different locations at the maximum rate permitted by the Twitter API (dev.twitter.com) from August 2, 2010 to May 28, 2011: this gives a sample rather than a complete set of relevant tweets due to the rate limiting. A by-product of the rate limiting is that smaller countries have relatively higher proportions of matching tweets in the data set (e.g., perhaps we obtained 10% of matching tweets for the US but 90% for Hong Kong). The seven locations were chosen to be places where English was widely used: The UK and Ireland (abbreviated to UK for convenience), the USA, Australia (sometimes abbreviated to Aus below), the Republic of South Africa (RSA), New Zealand (NZ), India and Hong Kong. Whilst English is not the most spoken language in RSA, India or Hong Kong, it was expected to be widely enough used to give data to analyse.

For each location, a Twitter query was constructed to return all Tweets in English within a specified geographic area. According to the Twitter API, "The location is preferentially taking from the Geotagging API, but will fall back to their Twitter profile"[12]. Hence the result will capture people geolocated by their mobile device or their profile. With the large area queries used, these two seem likely to be interchangeable. An alternative would be to identify Twitter users' locations from text in their profiles (Naaman et al., 2011), but this would be error-prone on the large scale necessary for this study. The choice of countries was relatively arbitrary: it excluded Canada because its shape made it awkward to collect Tweets from it because the geolocation search pattern must be a circle (specified by radius and centre longitude and latitude). Ireland is also predominantly English-speaking but its Tweets were gathered by the UK query as the two fell within a convenient single circle. Hence, Canada is the main (predominantly) English-speaking nation omitted. The Hong Kong data was not

---

[2] https://dev.twitter.com/docs/api/1/get/search

used because too little was collected to identify meaningful trends – for example, all the top 10 Hong Kong terms (see below for extraction method) appeared to be generated by spam rather than genuine trends (email, bny, etc, club, flyer, charge, any1, card, design, contact). Non-circle shapes could have been built from multiple overlapping circles but this would have been inefficient use of the limited Twitter API query capability. A total of 440 million English tweets were collected from 15.8 million different accounts (in millions of tweets by country: USA 141, UK 134, Australia 67.1, India 55.0, RSA 26.3, NZ 12.0, Hong Kong 5.41).

For each country, the time series scanning method (Thelwall & Prabowo, 2007) was used to identify the top 1,000 spiking words. This method creates a daily time series for every word (after converting plural words to singular and excluding hashtags, which have an advantage with this approach) used in every Tweet by calculating the proportion of tweets from the location (e.g., UK, India) containing the word. For each time series, the maximum spike size calculated is the maximum increase in the proportion of tweets containing the word during a single day, subtract the average of all previous days. For example, a score of 0.1 for the word rapture indicates that on one day the proportion of tweets containing "rapture" was 0.1 higher than the average proportion of tweets containing the word over all previous days.

The 1,000 most spiking words were then extracted and used as the basis for identifying topics in Twitter in the six different locations. This method has been shown to be useful to identify news-related topics within periodic text data (Thelwall & Prabowo, 2007), including from Twitter (Mathioudakis & Koudas, 2010; Thelwall et al., 2011), because such topics are often associated with one or more terms that cause a large spike. Nevertheless there is some redundancy because the same topic can be associated with multiple spiking words (e.g., both "royal" and "wedding" might spike for the UK royal wedding topic). Other topic detection methods avoid this limitation (e.g., Cataldi, Caro, & Schifanella, 2010; Ding, 2011) or may use additional filtering steps (e.g., Alonso, Carson, Gerster, Ji, & Nabar, 2010; Becker, Naaman, & Gravano, 2011b), identify specific types of event (Becker, Naaman, & Gravano, 2011a), use text-based characteristics, such as sentiment and URL inclusion (Naveed, Gottron, Kunegis, & Alhadi, 2011) or automatically classify extracted events (Genc, Sakamoto, & Nickerson, 2011). Nevertheless, the advantage of time series scanning method used is that it is (a) simple and transparent so that it can be corrected with additional human input, as described below, and (b) able to be combined with keyword frequencies as a way to identify the most important topics over a given time period. These characteristics are also true of another method used (Naaman et al., 2011), which was not chosen because its inverse document frequency component disadvantages relatively high frequency words that might nevertheless be relevant in trends, such as "Obama".

The top 1000 terms for each of the six countries (excluding Hong Kong) gave a final data set for analysis of 117.8 million tweets. To give a more specific idea of the numbers involved, the 1000th term for New Zealand (i.e., the least significant keyword in the country used with the fewest tweets) was "fergie", which occurred in 601 New Zealand tweets. In contrast, the top term for the US was "thanksgiving", which occurred in 76,973 US tweets.

Analysis

To calculate the overall trending topic similarity between countries, the top 50 spiking terms from the basic set of 1,000 described above were selected for each country as a basis for comparison. The number 50 was chosen because ranks below 50 seemed to be less reliable for the smallest data set, New Zealand (the 50th word for New Zealand, "grammy", occurred in 2,348 New Zealand tweets). This was a qualitative judgement from observation of the data rather than the result of a mathematical comparison. This judgment needs to be qualitative because the large number of words processed in the data set (11,772,149 unique terms), make it statistically inevitable that quite large spikes occur at random. Each of the top 50 words for each country was manually checked to see whether it referred to the same story as a higher ranked word and removed from the list if it did. Two words were considered to represent the same topic if they described the same event occurring on the same day or a small event associated with a large event on another day. As an example of the latter, a memorial event or silence for a tragedy was coded as referring to the same event as the tragedy even though it occurred on a different day. This rule was used to stop one event from dominating the results and

hence undermining the statistical value of tests run on the resultant data (due to violating the independence condition).

Each term in each filtered top 50 list was then checked against the top 1000 list of all other countries and the rank of the same term was recorded if (a) the term was in the top 1,000 and (b) the topic associated with the word was the same. Words failing this test were assigned a nominal rank of 1001 to indicate that they were unranked for a statistical comparison. A Spearman correlation was calculated between the rank list of the filtered top 50 terms of each country and the ranks of the terms for the other countries, obtained as described above. For each country this gives a measure of the extent to which the news topics that it considers important are also considered important by each other country.

To identify the main types of topics present in the data, the filtered top 50 trending topics for each country were categorised by the second author using an inductive approach to identify the main topics represented and to construct a short description of the categories. Each of the automatically-identified "trending topics" was essentially a keyword together with a list of tweets from the country containing the keyword, together with a time-series graph of the relative frequency of matching tweets associated with the country concerned. For instance, the keyword "independence" within the India data set was identified as being caused by Independence Day celebrations in India. The purpose of the subsequent categorisation process was to identify the broad type of reason causing each keyword to trend. These reasons were discovered inductively by examining all the terms and devising categories that matched common themes. For instance, in the above example the Indian Independence Day trending topic, amongst others, gave rise to the broad category of "Festival or religious".

The scheme (keywords and short descriptions of associated events, see Table 1) was then given to two independent people to code using the category descriptions. This non-standard approach was used because the categories were fixed and so it would not be appropriate to train coders on a random sample, as would normally be standard practice (Neuendorf, 2002). Cohen's Kappa (Cohen, 1960; Neuendorf, 2002) agreement rates for this were 'excellent' (0.820, 0.814. 0.840), giving some confidence in the results (Fleiss, 1981). The results reported are based upon the majority decision for each type, with some exceptions: the death of Osama Bin Laden was categorised as politics rather than human interest; the death of Indian Guru Sri Sathya Sai Baba was categorised as festival or religious rather than human interest, as was earth hour. These anomalies were caused by the weak initial category descriptions. There were also some overlaps between categories that caused discrepancies. For example, some weather-based disasters could be categorised as weather or human interest (the latter was chosen in all cases) and some events could be human interest or media (e.g., judgement day cult prediction, UK royal wedding and some could be politics or media (e.g., a political documentary, and TV interview with a former politician).

Results

The top 50 trending topics collected from the six different countries, after duplicate topic elimination, were analysed to identify the typical types of topic found and to identify international differences.

Main topics of interest

The most common trending topic was a festival or religious occasion (Table 1). The messages for these were often simple variants of "Happy […] day" and spiked on the day of the event. The topics included both religious events (e.g., Easter) and secular events (e.g., Independence Day). This type of topic was particularly numerous for India, for both secular and religious events.

The second most common type of topic was media, particularly for South Africa and New Zealand. This broad category includes TV shows, such as New Zealand's *Outrageous Fortune*, as well as U2 pop concerts and film premieres. This category included events, like the UK royal wedding, that could be conceived as political (a future head of state getting married) or human interest (a wedding) but would probably be viewed primarily as a TV show for most interested people.

Political events, such as elections, budgets, and protests formed the third most common topic. The political events include some of primarily national significance, such as budgets and elections, as well as some of international importance, such as the death of Osama Bin Laden and a London court

appearance for Wikileaks' Julian Assange. Some political events were national in scope, such as protests against Mubarak in Egypt, but generated international interest.

Human interest stories mainly involved disasters, such as the Japanese Tsunami, or individual deaths, including those of media figures. This category also included school exams, although these may not be often thought of as a human interest news topic.

Sports stories created significant national differences, even for international competitions. Only two sports were included for multiple countries: cricket (UK, India, South Africa) and soccer (UK, New Zealand, South Africa). The sports mentioned in only one country were basketball (US), American football (US), horse racing (Australia), Australian rules football (Australia), tennis (UK), rugby (South Africa), and athletics (India, hosting the Commonwealth Games). Cricket may have a high profile because it is dominated by English speaking countries and the Cricket World Cup occurred during the period surveyed.

Table 1. Topic type for the top 50 trending topics in each country, after eliminating duplicate topics. Note that the human interest category excludes deaths of political or religious figures.

		Number of topics					
Topic type	Definition of topic type	UK	NZ	Aus	India	USA	RSA
Festival or religious	Festival or religious event and associated activities, even if unofficial	13	9	11	15	12	12
Media	TV show, film, concert or associated events	4	6	4	1	3	7
Political events	Political events; elections; budgets; protests; rebellions	3	2	3	8	4	3
Human interest	Disasters; death or injury of named individuals; exams (other mass participation events)	4	6	5	1	4	4
Sport	Sporting event or associated activities	3	1	2	6	3	4
Natural events	Weather, seasons or astronomy - but not with a human interest angle	2	2	1	0	1	1
Viral text	Viral texts, including jokes	2	2	1	0	1	1
Technology	Technology developments or companies	0	0	1	1	0	0
Total		31	28	28	32	28	32

Similarity between countries for top trending topics

In terms of the types of trending topics, India stands out for having many festivals and political events but few media events, human interest stories and natural events (Table 1). The low representation of media events may be due to India having a strong home-grown Hindi media industry, such as Bollywood for films, and so the natural language for discussing related events may be Hindi rather than English. Bollywood has an international sphere of influence in Indian diaspora communities (Alessandrini, 2001; Bhattacharya, 2004) that may have English as a first language, but this was not enough to be represented in the Twitter results for any country. The only media event present in for India was the UK royal wedding, and this had the lowest rank of any country. India alone had two highly ranked national political issues: the anti-corruption protests of Anna Hazare and a court verdict relating to a Hindu/Muslim religious dispute in Ayodhya.

Table 2 reports the correlations between the rank order of the top up to 50 trending topics, after excluding duplicate topics in the top 50, with the rank order of the same topics in each other country. The table is in descending order of mean correlation (final column). From these national means, topics of most interest to the US tend to also be found interesting by other countries, in line with previous findings about the dominance of the US for international news. Topics of most interest to India attract by far the least interest in other countries. In the latter case, this is partly because many of the main topics of interest are associated with the Hindu religion, which is predominantly found in India, and festival days that are only found in India, such as Children's Day.

The six countries show much more similar average levels of interest in other countries' topics (the last row of Table 2) but the order is approximately reversed from the above: The US is least

interested in other countries' topics and India is most interested in other countries' topics. The latter is unexpected, given the cultural differences between India and the other countries. It is primarily due to the averaging process used because other countries include in their averages a low score for India.

Table 2. Spearman correlations between the rank order of the top up to 50 trending topics in each country (rows) with the same topics in other countries (columns). For example the rank correlation of the top US trending topics with the same topics in the UK is 0.69.

Source\Other	USA	UK	NZ	RSA	Aus.	India	*Mean*
USA	-	0.69	0.62	0.63	0.57	0.70	0.64
UK	0.55	-	0.51	0.58	0.54	0.59	0.56
NZ	0.42	0.48	-	0.43	0.49	0.55	0.47
RSA	0.41	0.49	0.49	-	0.50	0.41	0.46
Aus.	0.44	0.50	0.56	0.29	-	0.34	0.43
India	0.15	0.13	0.22	0.21	0.22	-	0.18
Mean	0.39	0.46	0.48	0.43	0.46	0.52	-

Limitations

An important limitation of the study from the perspective of attempting to draw conclusions beyond Twitter is that the Twitter data is not representative of the general population of the countries surveyed. The use of Twitter for sampling is likely to cause a bias towards more technical users and hence towards technological issues, although only two technology topics were found in the top 50s investigated so the impact of this may not be large. This bias is an important consideration for the objective of using Twitter to gain insights into international differences in news coverage.

From the perspective of Twitter users themselves (i.e., ignoring non-users), probably the most significant limitation is that the Twitter data from India is likely to represent a different segment of the population, and even a different segment of the Twitter using population, than in the other countries. This is partly because India is a poorer country and frequent internet access (e.g., via a mobile device) is probably limited to a smaller percentage of the population and the wealthiest people in the country. Tweeting in English in India also suggests a high level of fluency in English as a second language and a motivation for using it. These are likely to cause biases towards topics of concern to educated people and perhaps also towards people with an international orientation or even a special attraction to the US as a world-leading nation. This seriously undermines the differences found for India. Moreover, the findings should not be extrapolated to refer to the whole population of India.

The method of eliminating terms referring to events associated with a higher ranking word caused problems with the cricket world cup. This consisted of a series of discrete events (matches) relating to different pairs of countries and on different days, but connected due to being part of a common competition. These matches were coded as separate events rather than as one single "cricket world cup" topic because of the clearly different dates and national interests for the different matches. This decision nevertheless had an impact on the results and made this single long event important for India in particular.

Discussion

In response to the main research questions, the main types of trending topics for these countries (Q1) are: festivals or religious events, media events, politics, human interest and sport. The types of topics vary by country (Q2) both in terms of differences in the types of topics discussed and in terms of the specific topics discussed. In particular, H2a is confirmed: Trending topics of interest to the USA are of most interest to other countries and trending topics of interest to India are of least interest to other countries. In contrast, H2b is only partly confirmed: The USA is least interested in trending topics of interest to other countries, but India is not also the least interested in trending topics of interest to other countries. Even though Indian Twitter users seem to focus on topics that are mainly of interest within India, they maintain an awareness of international topics too. This last finding may well reflect the demographic of the Indian twitter users captured by the data rather than the typical Indian citizen.

Perhaps the strangest fact is that the trending topic thanksgiving, which refers to an exclusively US event with no significance in the rest of the world (except Canada, but on a different day), is in the top 50 for all countries except India, where it has rank 57. In conjunction with H2a and the confirmed part of H2b, his seems to demonstrate the unique power of the US to attract interest from the rest of the world, even at the informal level of tweeting and for non-news events. Examples of tweets from New Zealand include, "to celebrate American thanksgiving I ate sushi", "happy thanksgiving to all of my American friends", and just "happy thanksgiving" as well as some from people apparently celebrating thanksgiving (e.g., "need a way to work off this thanksgiving gut"), perhaps Americans abroad. A possible implication of this is that international biases in interest are so deeply ingrained that either they would continue to some extent without media agenda setting or that the public have learned biases from experience with the media over time and accepted them, presumably subconsciously for the majority of people. Alternative explanations are possible, however, such as spam or learning about Thanksgiving from US television shows or movies, given their widespread popularity abroad (Bielby & Harrington, 2008). Although such shows and movies would presumably not synchronise abroad with Thanksgiving Day, they might create an awareness of the festival that would translate into interest on the appropriate day. In the latter case this suggests a reasonably long-lasting influence for non-news media. This should not be surprising since in various contexts US television and movies have been shown to be able to create awareness of issues and long term change, both for good (e.g., Winsten, 2000) and bad (e.g., Viswanath, Blendon & Vallone, 2008).

Conclusions

Although Twitter seems to be primarily used by people to report about themselves, analysing trending topics can give insights into news-related issues that are widely discussed by the minority of users that focus on external events. In all the countries investigated, the trending topics identified covered a wide variety of types (primarily festivals or religious events, media events, politics, human interest and sport) although there were some differences. International influence in Twitter seems to broadly follow similar patterns to those of news reporting, which is not surprising because news coverage presumably impacts what people tweet about. In particular, topics of interest to the US seem to be typically picked up in other nations, including exclusively US holidays and US national sport events. This suggests that some overseas populations are genuinely interested in US news, which they presumably find out about from mainly national news sources. Conversely, the other countries surveyed were relatively uninterested in Indian issues. This may be partly due to cultural differences, low coverage of Indian topics in the mainstream national media or a perception that India is economically or socially unimportant and therefore not worth taking much notice of. Perhaps most surprising is the interest in a domestic US issue: Thanksgiving Day, which was presumably little covered in the international media. This suggests a much deeper embedding of international imbalances within individuals that goes further than merely accepting news media agendas.

An implication of this study from a methodological perspective is that it is possible to compare the main trending topics of interest between nations in order to gain insights into international commonalities and differences, although the results cannot be easily generalised beyond Twitter users to all people in each country. The method used is restricted to countries in which a significant number of people tweet in English or another shared language and the power of the findings is limited by differences in the extent to which the chosen language is used in any country. The method is also limited to the top trending topics rather than the typical discussion topics, although the latter could be easily investigated by taking a random sample from the different national locations and conducting a content analysis of them. Based on the literature review, this seems likely to give results dominated by "me now" tweets rather than general interest topics. A new method would be needed to identify topics of more general interest to compare between nations – for example by only sampling retweets. It would also be possible to extend the basic method to a multi-lingual version by removing the single language restriction but this may weaken any findings produced by introducing linguistic rather than topic-based differences between nations.

More generally, the method can be extended to compare specific topics of interest between countries using Twitter. For instance, to focus on a specific topic, such as a given health issue, tweets

could be collected or filtered using a set of relevant queries and then the results compared between countries. This comparison could be based upon trending subtopics, as in the current paper, or a content analysis of a random sample of all relevant tweets, as discussed in the paragraph above. This would give an unobtrusive and relatively quick way to conduct international comparisons, although biased by the restriction to Twitter users.

Finally, an implication of the results for new media issues is confirmation from a new source (Twitter, albeit with its coverage/user group biases) that the current internationally imbalanced news media coverage does not seem be out of step with public news interests, whether it leads them or not. Whilst this should be reassuring for news media organisations, the apparent deep embedding of international differences of importance discussed above – for example, to the extent that the US-specific Thanksgiving Day is a major discussion topic in most countries surveyed – is likely to be viewed as a problem by those that believe that the media are too powerful or too easily led by politicians, public relations teams or powerful interest groups (e.g., Bennett, Lawrence, & Livingston, 2007). This apparent international public acceptance of imbalances is likely to be seen as even more of a problem by those that take the stronger position that powerful interests are often successful at using the media to manufacture consent for their goals, including government foreign policy goals (Herman & Chomsky, 1988).

Acknowledgement

Appendix

Table 3. Top topics for the UK and word rank for the same topic in other countries. *The word is primarily associated with another topic for this country.

Word	Rank of topic in country						UK topic
	UK	NZ	Aus	India	USA	RSA	
valentine	1	6	2	4	2	2	Valentine's day Feb. 14, 2011
easter	2	7	1	25	4	10	Easter day, Apr. 24, 2011
snow	3	199*	279	677	32*	171	Snow (UK), Nov. 31, 2011
wedding	4	2	3	20	15	3	UK Royal Wedding Apr. 29, 2011
bin	5	9	7	9	5	8	Osama Bin Laden death May 2, 2011
pancake	8	898	183	-	288	260	Pancake day Mar 8, 2011
mother	9	16	5	8	3	1	Mother's day 8 May 2011
cock	10	173	187	213*	210	238	Viral: #puttingcockinsongname Apr. 19, 2011
halloween	11	23	17	50	8	39	Halloween Oct. 31, 2010
japan	13	3	14	13	9	13	Earthquake and Tsunami in Japan Mar. 11, 2011
miner	14	14	44	118	26	15	Chilean miners released, Oct. 13, 2010
fool	16	43	28	31	21	17	April Fool's day April 1, 2011
bank	17	-	-	-	-	-	Bank Holiday (UK), Aug. 31, 2010
moon	18	110	147	80	144	130	Full moon (biggest of the year) Mar. 20, 2011
st	19	38	-	481	17	31	St Patrick's day Mar. 17, 2011
christmas	20	32	20	281	37	121	Christmas
vote	23	-	4*	141*	13*	5*	UK Alternative Vote referendum, May 5, 2011
exam	26	-	-	-	-	94*	Exam results UK schools May 16, 2011
thanksgiving	27	28	23	57	1	38	Thanksgiving (US) Nov. 26 2010
murray	28	537	93	507	479	577	Andy Murray (UK) Aus. tennis final, Jan. 30, 2011
matt	29	-	-	913	-	457	X-factor final (UK TV show), Dec 12, 2011
pope	31	933	346	945	522*	394	Pope UK visit, Sept. 16, 2011
student	32	-	-	-	-	-	UK tuition fees student demo Nov. 11, 2011
cup	33	-	22*	11*		-	Soccer world cup 2018 venue decided Dec. 2, 2010
firework	36	154	241*	-	470	341	Guy Fawkes Night (UK) Nov 5, 2010
delay	37	-	-	-	-	-	Multiple minor UK motorway delays Feb. 21, 2011
cheryl	38	-	-	-	-	96	X-factor (UK TV show) Sept 29, 2010
torre	39	709	321	234	578	85	Torres' UK soccer club switch Feb. 1, 2011
cher	41	-	294	-	867	46	X-factor (UK) final, Dec. 11, 2010
pant	43	19	13	212	24	89	Viral: #moviesinmypants Oct. 31, 2009
children	44	123	-	34	-	-	Children in Need UK TV fundraising Nov. 19, 2010
rapture	48	29	39	367	35	127	Judgement day cult prediction (US) May 21, 2011
gcse	49	-	-	-	-	-	Exam results UK schools Aug, 24, 2010
eurovision	50	-	135	-	742	-	Eurovision song contest May 15, 2011

Table 4. Top topics for New Zealand and word rank for the same topic in other countries. *The word is primarily associated with another topic for this country.

	Rank of topic in country						
Word	**NZ**	**UK**	**Aus**	**India**	**USA**	**RSA**	**NZ topic**
christchurch	1	76	16	735	219	348	Earthquake in Christchurch, NZ Feb. 22, 2011
wedding	2	4	3	20	15	3	UK Royal Wedding Apr. 29, 2011
japan	3	13	14	13	9	13	Earthquake and Tsunami in Japan Mar. 11, 2011
valentine	6	1	2	4	2	2	Valentine's day Feb. 14, 2011
easter	7	2	1	25	4	10	Easter day, Apr. 24, 2011
bin	9	5	7	9	5	8	Osama Bin Laden death May 2, 2011
miner	14	14*	44*	118*	26*	15*	Pike River mining tragedy (NZ) Nov 24, 2010
tornado	15	-	524*	-	112*	-	Tornado in Albany, NZ, May 3, 2011
mother	16	9	5	8	3	1	Mother's day 8 May 2011
storm	18	379*	52*	-	155*	-	Storm in Wellington, NZ Sept. 17, 2010
pant	19	43	13	212	24	89	Viral: #moviesinmypants Oct 31, 2010
halloween	23	11	17	50	8	39	Halloween Oct 31, 2010
aftershock	25	-	-	-	-	-	Christchurch earthquake aftershock Sept. 4, 2010
thanksgiving	27	27	23	57	1	38	Thanksgiving (US) Nov 26 2010
rapture	28	48	39	367	35	127	Judgement day cult prediction (US) May 21, 2011
henry	30	662*	-	-	-	-	Paul Henry (NZ) ridicule of Indian diplomat Sheila Dikshit Oct. 5, 2010
christmas	31	20	20	281	37	121	Christmas
memorial	33	-	891	-	188	-	Christchurch earthquake memorial service March 17, 2011
st	37	19	-	481	17	31	St Patrick's day Mar. 17, 2011
hobbit	39	-	-	-	-	-	Meeting about filming The Hobbit in NZ Oct. 27, 2010
fool	40	16	28	31	21	17	April Fool's day Apr. 1, 2011
grammy	42	65	47	117	22	37	Grammy awards Feb. 14, 2011
fortune	43	-	-	-	-	-	Outrageous Fortune TV series finale - NZ Nov. 9, 2010
budget	44	113*	69*	42*	415*	103*	NZ budget May 19, 2011
u2	46	-	406*	-	-	24*	U2 concert in NZ Nov. 25, 2011
cyclone	48	370	6	704*	411	571	Queensland floods and Cyclone Yasi (Aus) Feb. 2, 2011
lube	50	122	190	-	314	213	Viral: #changelovetolubesong Mar. 9, 2011

Table 5. Top topics for Australia and word rank for the same topic in other countries. *The word is primarily associated with another topic for this country.

Word	Rank of topic in country						Australia topic
	Aus	UK	NZ	India	USA	RSA	
easter	1	2	7	25	4	10	Easter day, Apr. 24, 2011
valentine	2	1	6	4	2	2	Valentine's day Feb. 14, 2011
wedding	3	4	2	20	15	3	UK Royal Wedding Apr. 29, 2011
vote	4	23*	-	141*	13*	5*	Federal Election (Australia) Aug. 21, 2010
mother	5	9	16	8	3	1	Mother's day 8 May 2011
cyclone	6	370	48	704*	411	571	Cyclone Yasi (Australia) Feb. 2, 2011
bin	7	5	9	9	5	8	Osama Bin Laden death May 2, 2011
flood	10	693	213	-	670	346	Queensland (Australia) floods Jan. 13, 2011
father	11	-	58	-	-	-	Father's Day Sept. 5, 2010
pant	13	43	19	212	24	89	Viral: #moviesinmypants Oct. 31, 2009
japan	14	13	3	13	9	13	Earthquake and Tsunami Japan Mar. 11, 2011
christchurch	16	76	1	735	219	348	Earthquake in Christchurch, NZ Feb. 22, 2012
halloween	17	11	23	50	8	39	Halloween Oct. 31, 2010
christmas	20	20	31	281	37	121	Christmas
cup	22	33*	-	11*	-	-	Melbourne Cup horse race (Australia) Nov. 2, 2011
thanksgiving	23	27	27	57	1	38	Thanksgiving (US) Nov. 25, 2010
howard	25	603	-	-	850*	-	Ex Aus. PM John Howard on Aus. TV Oct. 25, 2010
anzac	27	-	69	-	-	-	Anzac Day (Australia-New Zealand) Apr. 25, 2011
fool	28	16	40	31	21	17	April Fool's day Apr. 1, 2011
afl	35	-	-	-	-	-	AFL Grand Final (Australia) Australian rules football, Sept. 25, 2010
assange	38	138	287	101	165	299	Wikileaks' Assange denied rape bail in London (UK), Dec .7, 2010
rapture	39	48	28	367	35	127	Judgement day cult prediction (US) May 21, 2011
spring	41	450*	73	-	124*	7	1st day of Spring tomorrow, Aug 31, 2010
miner	44	14	14*	118	26	15	Chilean miners released, Oct. 13, 2010
oprah	45	-	-	-	377*	558*	Oprah Winfrey records show in Sydney, Australia, Dec. 14, 2010
earth	46	-	197	96	61*	49	Earth hour, Mar. 26, 2011
grammy	47	65	42	117	22	37	Grammy awards Feb. 14, 2011
apple	50	-	-	-	-	-	Apple online event in US, Sept. 1, 2010

Table 6. Top topics for India and word rank for the same topic in other countries. *The word is primarily associated with another topic for this country.

Word	Rank of topic in country						India topic
	India	UK	NZ	Aus	USA	RSA	
osama	1	12	10	9	7	4	Osama Bin Laden death May 2, 2011
independence	2	449*	-	-	-	-	Indian Independence Day Aug. 15, 2010
diwali	3	-	883	-	-	866	Diwali Hindu festival Nov. 4, 2010
valentine	4	1	6	2	2	2	Valentine's day Feb. 14, 2011
republic	5	-	-	-	-	-	Indian Republic Day Jan. 26, 2011
anna	6	-	-	222*	-	-	Anna Hazare (India) anti-corruption protest Apr. 7, 2011
verdict	7	-	-	488	-	-	Allahabad (India) High Court verdict on Ayodhya dispute Sept. 30, 2010
mother	8	9	16	5	3	1	Mother's day 8 May 2011
cup	11	33*	-	22*	-	-	Cricket world cup final, India v Sri Lanka Apr. 2, 2011
japan	13	13	3	14	9	13	Earthquake and Tsunami in Japan Mar. 11, 2011
match	14	-	-	-	-	-	Cricket world cup semi-finals, India v Pakistan Mar. 30, 2011
holi	15	-	-	-	-	-	Holi Hindu Spring festival (colours), Mar. 20, 2011
opening	16	-	-	150	222*	322	Commonwealth Games opening ceremony Delhi (India) Oct. 3, 2010
obama	17	54*	26*	36*	14*	32*	Obama Delhi (India) visit Nov. 7, 2010
royal	18	6	11	8	20	6	UK Royal Wedding Apr. 29, 2011
ganesh	19	-	-	-	-	-	Ganesh Chaturthi Hindu festival, Sept. 11, 2011
eid	24	184	-	-	709	174	Muslim festival Eid, Sept. 11, 2011
easter	25	2	7	1	4	10	Easter day, Apr. 24, 2011
teacher	27	-	-	-	-	-	Teacher's Day (india), Sept. 5, 2010
fool	31	16	40	28	21	17	April Fool's Day Apr. 1, 2012
children	34	44*	-	-	-	-	Children's Day (India) Nov. 14, 2010
ipl	35	-	-	-	-	-	Start of Indian Premier League cricket, Apr. 8, 2011
ireland	37	-	758*	554*	-	136*	Cricket world cup, India v Ireland, March 7, 2011
sehwag	39	-	-	-	-	-	India vs. Bangladesh cricket world cup opener Feb. 19, 2011
laxman	40	-	-	-	-	-	Australia-India cricket test match Oct 5, 2010
sai	41	-	-	-	-	-	Death of Indian Guru Sri Sathya Sai Baba Apr. 24, 2011 (Hindu/Muslim mix)
budget	42	113*	44*	69*	415*	103*	Indian Budget, Feb. 28, 2011
ponting	43	-	-	244*	-	805	Cricket world cup Australia vs. India Mar. 24, 2011
onam	44	-	-	-	-	-	Hindu festival, Kerala, India Aug. 23, 2010

pm	45	-	-	-	-	-	Indian Prime Minister press conference Feb. 16, 2011
bihar	48	-	-	-	-	-	Indian elections Nov 24, 2011
nokia	49	352	261	267	274	338	Nokia-Microsoft partnership Feb. 11, 2011
halloween	50	11	23	17	8	39	Halloween Oct. 31, 2010

Table 7. Top topics for the USA and word rank for the same topic in other countries. *The word is primarily associated with another topic for this country.

Word	Rank of topic in country						USA topic
	USA	UK	NZ	Aus	India	RSA	
thanksgiving	1	27	27	23	57	38	Thanksgiving (US) Nov. 25, 2010
valentine	2	1	6	2	4	2	Valentine's day Feb. 14, 2011
mother	3	9	16	5	8	1	Mother's day 8 May 2011
easter	4	2	7	1	25	10	Easter day, Apr. 24, 2011
bin	5	5	9	7	9	8	Osama Bin Laden death May 2, 2011
halloween	8	11	23	17	50	39	Halloween Oct. 31, 2010
japan	9	13	3	14	13	13	Earthquake and Tsunami in Japan Mar. 11, 2011
vote	13	23*	-	4*	141*	5*	US Senate elections, Nov. 2, 2010
wedding	15	4	2	3	20	3	UK Royal Wedding Apr. 29, 2011
st	17	19	37	-	481	31	St Patrick's Day Mar. 17, 2011
veteran	19	-	538	734	952	-	Veteran's Day (US) Nov. 12, 2010
fool	21	16	40	28	31	17	April Fool's day Apr. 1, 2012
grammy	22	65	42	47	117	37	Grammy awards Feb. 14, 2011
pant	24	43	19	13	212	89	Viral: #moviesinmypants Oct. 31, 2009
nate	25	63	51	128	454	21	Death of Nate Dogg (US) Mar. 16, 2011
miner	26	14	14*	44	118	15	Chilean miners released, Oct. 13, 2010
labor	28	965	-	82*	889	-	Labor Day (US), Sept. 6, 2010
super	29	-	-	-	-	-	Super bowl (US) Feb. 7, 2011
snow	32	3*	195*	279*	677*	171*	Snowstorm in Midwest, US Feb. 2, 2011
rapture	35	48	28	39	367	127	Judgement day cult prediction (US) May 21, 2011
knick	36	-	842*	-	828	460	Knicks vs. Celtics basketball (US), Apr. 18, 2011
christmas	37	20	31	20	281	121	Christmas
harry	40	42*	67*	59*	179	74*	Release of Harry Potter and the Deathly Hallows, Part I, Nov. 19, 2010
egypt	41	136*	146	72	68*	68	Internet shut in Egypt in response to unrest, Jan. 28, 2011
purple	43	177	208	142	-	357	Spirit Day (international) -wear purple to support LGBT bullying victims Oct. 20, 2010
charlie	44	250	172	81	696*	317	Charlie Sheen's twins taken away Mar. 2, 2011
laker	49	-	225	675	686	729	Lakers vs. Dallas baketball (US), May 8, 2011

Table 8. Top topics for South Africa and word rank for the same topic in other countries. *The word is primarily associated with another topic for this country.

| Word | Rank of topic in country | | | | | | RSA topic |
	RSA	UK	NZ	Aus	India	USA	
mother	1	9	16	5	8	3	Mother's day 8 May 2011
valentine	2	1	6	2	4	2	Valentine's day Feb. 14, 2011
wedding	3	4	2	3	20	15	UK Royal Wedding Apr. 29, 2011
osama	4	12	10	9	1	7	Death of Osama Bin Laden May 2, 2011
vote	5	23*	-	4*	141*	13*	South African Elections, May 18, 2011
spring	7	450*	73	41		124*	1st day of Spring tomorrow, Sept 1, 2010
easter	10	2	7	1	25	4	Easter day, Apr. 24, 2011
japan	13	13	3	14	13	9	Earthquake and Tsunami in Japan Mar. 11, 2011
cleavage	14	915	-	-	-	-	Wonderbra's National Cleavage Day RSA Apr. 15, 2011
miner	15	14	14*	44	118	26	Chilean miners released, Oct. 13, 2010
fool	17	16	40	28	31	21	April Fool's day Apr. 1, 2012
freedom	18	-	-	-	102*	394*	Freedom Day (South Africa) April 27, 2011
aid	19	255	629	320	146	104	World AIDS day, Dec 1, 2010
nate	21	63	51	128	454	25	Death of Nate Dogg (US) Mar. 16, 2011
u2	24	-	46	406	-	-	U2 concert South Africa Feb 13, 2011
shark	25	-	-	-	-	-	Sharks vs. Bulls, rugby South Africa Oct 16, 2011
arsenal	27	90	-	235	262	769	Arsenal vs. Barcelona, European soccer, Mar. 8, 2011
st	31	19	37	-	481	17	St Patrick's Day Mar. 17, 2011
mandela	33	-	-	-	-	-	Nelson Mandela (RSA) hospitalised, Jan. 27, 2011
vagina	35	-	-	-	-	-	#replacehousewithvagina, Mar 30, 2011
grammy	37	65	42	47	117	22	Grammy awards Feb. 14, 2011
thanksgiving	38	27	27	23	57	1	Thanksgiving (US) Nov. 25, 2010
halloween	39	11	23	17	50	8	Halloween Oct. 31, 2010
braai	40	-	-	-	-	-	Unofficial National Braai (BBQ) Day RSA Sept 24, 2010
proteas	41	-	-	-	-	-	Cricket South Africa vs. New Zealand, Mar 25, 2011
pirate	42	622	-	413*	761	818	Pirates of the Caribbean: On stranger tides, released, May 21, 2011
chief	44	-	972*	-	-	581*	Orlando Pirates vs. Kaiser Chiefs (RSA soccer) Nov 13, 2011
cher	46	41	-	294	-	867	X-factor (UK) final, Dec 11, 2010
khanyi	47	-	-	-	-	-	Program critical of wealthy RSA socialites Nov 30, 2011
sama	48	-	-	-	-	-	South African Music Awards, May 21, 2011
earth	49	-	197	46	96	61*	Earth hour, Mar. 26, 2011
mubarak	50	144	228	91	62	62	Hosni Mubarak under pressure to quit in Egypt, Feb. 11, 2011

References

Abel, F., Gao, Q., Houben, G.-J., & Tao, K. (2011). Analyzing user modeling on Twitter for personalized news recommendations. *Lecture Notes in Computer Science*, 6787/2011, 1-12.

Ahlers, D. (2006). News consumption and the new electronic media. *Harvard International Journal Of Press-Politics, 11*(1), 29-52.

Alessandrini, A. C. (2001). "My heart's Indian for all that": Bollywood film between home and diaspora. *Diaspora: A Journal of Transnational Studies, 10*(3), 315-340.

Alexa (2011). Top sites. Retrieved December 6, 2011 from: http://www.alexa.com/topsites.

Alonso, O., Carson, C., Gerster, D., Ji, X., & Nabar, S. U. (2010). Detecting uninteresting content in text streams. In *Proceedings of the SIGIR 2010 Workshop on Crowdsourcing for Search Evaluation (CSE 2010)* (pp. 39-42).

Becker, H., Naaman, M., & Gravano, L. (2011a). Beyond trending topics: Real-world event identification on Twitter. In N. Nicolov & J. G. Shanahan (Eds.), *Proceedings of the 5th International Conference on Weblogs and Social Media (ICWSM 2011)* (pp. Retrieved December 6, 2011 from: http://sm.rutgers.edu/pubs/becker2035-icwsm2011.pdf). Menlo Park, CA: The AAAI Press.

Becker, H., Naaman, M., & Gravano, L. (2011b). Selecting quality Twitter content for events. In N. Nicolov & J. G. Shanahan (Eds.), *Proceedings of the 5th International Conference on Weblogs and Social Media (ICWSM 2011)* (pp. Retrieved December 6, 2011 from: http://sm.rutgers.edu/pubs/becker2082-icwsm2011.pdf). Menlo Park, CA: The AAAI Press.

Bennett, W.L., Lawrence, R.G., & Livingston, S. (2007). *When the press fails: political power and the news media from Iraq to Katrina*, Chicago: University of Chicago Press.

Berger, G. (2009). How the internet impacts on international news: Exploring paradoxes of the most global medium in a time of 'hyperlocalism'. *International Communication Gazette, 71*(5), 355-371.

Besova, A. A., & Cooley, S. C. (2009). Foreign news and public opinion: Attribute agenda-setting theory revisited. *Ecquid Novi: African Journalism Studies, 30*(2), 219-242.

Bhattacharya, N. (2004). A 'basement' cinephilia. Indian diaspora women watch Bollywood *South Asian Popular Culture, 2*(2), 161-183.

Bielby, D, B. & Harrington, C.L. (2008). *Global TV: Exporting television and culture in the world market*, New York: New York University Press.

Blake, K.D., Viswanath, K., Blendon, R.J., & Vallone, D. (2008). The role of reported tobacco-specific media exposure on adult attitudes towards proposed policies to limit the portrayal of smoking in movies. *Pediatrics*, 121(1), e108-e117.

Bruns, A. (in press). How long is a tweet? Mapping dynamic conversation networks on Twitter using Gawk and Gephi. *Information, Communication & Society*.

Cataldi, M., Caro, L. D., & Schifanella, C. (2010). Emerging topic detection on Twitter based on temporal and social terms evaluation In *Proceedings of the Tenth International Workshop on Multimedia Data Mining table of contents* (pp. A4). New York, NY: ACM Press.

Chang, T.-K., Shoemaker, P. J., & Brendlinger, N. (1987). Determinants of international news coverage in the U.S. media. *Communication Research, 14*(4), 396-414.

Chang, T.-K., (1998). All countries not created equal to be news: World system and international communication, *Communication Research*, 25(5), 528-566.

Chen, G. M. (2011). Tweet this: A uses and gratifications perspective on how active Twitter use gratifies a need to connect with others. *Computers in Human Behavior, 27*(2), 755-762.

Chomsky, N. (2000). *A new generation draws the line: Kosovo, East Timor and the standards of the West*. New York: Verso Books.

Cohen, B. (1963). *The press and foreign policy*. Princeton, NJ: Princeton University Press

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20*(1), 37-46.

Crystal, D. (1997). *English as a global language*. Cambridge: Cambridge University Press.

Dann, S. (2010). Twitter content classification. *First Monday, 15*(12), Retrieved December 6, 2011 from: http://frodo.lib.uic.edu/ojsjournals/index.php/fm/article/view/2745/2681.

Der Doran, J. (2009). Virtuous war: Mapping the military-industrial-media-entertainment network. New York, NY: Routledge.

Ding, Y. (2011). Community detection: Topological vs. topical. *Journal of Informetrics, 5*(4), 498-514.

Dotan, A., & Zaphiris, P. (2010). A cross-cultural analysis of Flickr users from Peru, Israel, Iran, Taiwan and the UK. *International Journal of Web Based Communities, 6*(3), 284-302.

Entman, R. M. (1991). Framing U.S. coverage of international news: Contrasts in narratives of the KAL and Iran Air incidents. *Journal of Communication, 41*(4), 6-27.

Entman, R. M. (1993). Framing: Toward clarification of a fractured paradigm. *Journal of Communication, 43*(4), 51-58.

Entman, R. M. (1994). Projections of power: Framing news, public opinion, and U.S. foreign policy. Chicago, IL: University of Chicago Press.

Eysenbach, G. (2011). Can tweets predict citations? Metrics of social impact based on Twitter and correlation with traditional metrics of scientific impact. *Journal of Medical Internet Research, 13*(4), e123.

Farhi, P. (2009). The Twitter explosion American Journalism Review, 31(3), Retrieved March 20, 2012 from: http://www.ajr.org/article.asp?id=4756.

Fleiss, J. L. (1981). *Statistical methods for rates and proportions (2nd ed.).* New York: John Wiley.

Fujisaka, T., Lee, R., & Sumiya, K. (2010). Detection of unusually crowded places through micro-blogging sites. In *24th International Conference on Advanced Information Networking and Applications Workshops (WAINA 2010)* (pp. 467-472). Los Alamitos, CA: IEEE.

Genc, Y., Sakamoto, Y., & Nickerson, J. V. (2011). Discovering context: Classifying tweets through a semantic transform based on Wikipedia. *Lecture Notes in Computer Science, 6780/2011*, 484-492.

Golan, G. J. (2006). Assessing the influence of the New York Times on three network television evening news programs. *Journalism Studies*, 7(2), 323-333.

Golan, G. J. (2008). Where in the World Is Africa? *International Communication Gazette*, 70(1), 41-57.

Golder, S. A., & Macy, M. W. (2011). Diurnal and seasonal mood vary with work, sleep, and daylength across diverse cultures. *Science, 333*(6051), 1878-1881

Groshek, J. (2008). Homogenous agendas, disparate frames: CNN and CNN International coverage online. *Journal of Broadcasting & Electronic Media, 52*(1), 52-68.

Hale, S. A. (2012). Net increase? Cross-lingual linking in the blogosphere. *Journal of Computer-Mediated Communication*, 17(2), 135-151.

Heaivilin, N., Gerbert, B., Page, J. E., & Gibbs, J. L. (2011). Public health surveillance of dental pain via Twitter. *Journal of Dental Research, 90*(9), 1047-1051

Herman, E. S., & Chomsky, N. (1988). *Manufacturing consent: The political economy of the mass media*. New York: Pantheon Books.

Hermida, A. (2010). From TV to Twitter: How ambient news became ambient journalism. *M/C Journal*, 13(2), Retrieved March 20, 2012 from: http://journal.media-culture.org.au/index.php/mcjournal/article/viewArticle/2220.

Hernandez, B. A. (2011). Facebook reveals 2011′s most-popular status trends *Mashable.com*, Retrieved December 7, 2011 from: http://mashable.com/2011/2012/2007/facebook-reveals-2011s-most-popular-status-trends/.

Heverin, T., & Zach, L. (2012). Use of microblogging for collective sense-making during violent crises: A study of three campus shootings. *Journal of the American Society for Information Science and Technology*, 63(1), 34-47.

Honeycutt, C., & Herring, S. C. (2009). Beyond microblogging: Conversation and collaboration via Twitter. In *Proceedings of the 42nd Hawaii International Conference on System Sciences* (pp. 1-10). Los Alamitos, CA: IEEE.

Huang, H. (2010). Frame-rich, frame-poor: An investigation of the contingent effects of media frame diversity and individual differences on audience frame diversity, *International Journal of Public Opinion Research*, 22(1), 47-73.

Huang, J., Thornton, K. M., & Efthimiadis, E. N. (2010). Conversational tagging in twitter. In *Proceedings of the 21st ACM conference on Hypertext and hypermedia* (pp. 173-178). New York, NY: ACM Press.

Hughes, A. L., & Palen, L. (2009). Twitter adoption and use in mass convergence and emergency events. *International Journal of Emergency Management, 6*(3-4), 248-260.

Iyengar, S., Peters, M.D., & Kinder, D.R. (1982). Experimental demonstrations of the "not-so-minimal" consequences of television news programs. *American Political Science Review*, 76(4), 848-858.

Jackoway, A., Samet, H., & Sankaranarayanan, J. (2011). Identification of live news events using Twitter. In *Proceedings of the 3rd ACM SIGSPATIAL International Workshop on Location-Based Social Networks (LBSN '11)*. New York, NY: ACM Press.

Jansen, B. J., Sobel, K., & Cook, G. (2011). Classifying ecommerce information sharing behaviour by youths on social networking sites. *Journal of Information Science*, 37(2), 120-136.

Jansen, B. J., Zhang, M., Sobel, K., & Chowdury, A. (2009). Twitter power: Tweets as electronic word of mouth. *Journal of the American Society for Information Science & Technology, 60*(11), 2169-2188.

Java, A., Song, X., Finin, T., & Tseng, B. (2007). Why we Twitter: Understanding microblogging usage and communities. In *Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 workshop on Web mining and social network analysis* (pp. 56-65). New York, NY: ACM Press.

Kalb, M. (1994). A view from the press. In W. L. Bennett & D. L. Paletz (Eds.), *Taken by storm: Media, public opinion and U.S. foreign policy in the Gulf War* (pp. 3-6). Chicago: University of Chicago Press.

Kissau, K., & Hunger, U. (2008). Political online-participation of migrants in Germany. *German Policy Studies, 4*(4), 5-31.

Kolmer, C., & Semetko, H. A. (2010). International Television News. *Journalism Studies, 11*(5), 700-717.

Kwak, H., Lee, C., Park, H., & Moon, S. (2010). What is Twitter, a social network or a news media? In Proceedings of the 19th international conference on world wide web (pp. 591-600). New York, NY: ACM Press.

Kwak, N., Poor, N., & Skoric, M. M. (2006). Honey, I shrunk the world! The relation between Internet use and international engagement. *Mass Communication and Society*, 9(2), 189-213.

Larson, J.F. (1982). International affairs coverage on US evening network news, 1972-1979. In: W. Adams (Ed.), *Television's Coverage of International Affairs*, Norwood, NJ: Ablex.

Lerman, K., & Ghosh, R. (2010). Information contagion: An empirical study of the spread of news on Digg and Twitter social networks. In *Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media (ICSWM 2010)* (pp. 90-97). Menlo Park, CA: AAAI Press.

Lin, W.-Y., Lo, V.-h., & Wang, T.-L. (2011). Bias in television foreign news in China, Hong Kong, and Taiwan. *Chinese Journal of Communication 4*(3), 293-310.

Livingston, S. G. (1992). The politics of international agenda-setting: Reagan and North-South relations. *International Studies Quarterly*, 36(3), 313-330.

Loomis, K.D. (2009). A comparison of broadcast world news web pages: Al Jazeera English, BBC, CBS, and CNN, *Electronic News*, 3(3), 143-160.

Mathioudakis, M., & Koudas, N. (2010). TwitterMonitor: Trend detection over the Twitter stream. In *Proceedings of the 2010 International Conference on Management of Data (SIGMOD '10)* (pp. 1155-1157). New York, NY: ACM Press.

Matthes, J. (2009). What's in a frame? A content analysis of media framing studies in the world's leading communication journals, 1990-2005, *Journalism & Mass Communication Quarterly* 86(2), 349-367.

McCombs, M. E., & Bell, T. (1996). The agenda setting role of mass communication. In: M. Salwen & D. Stacks (eds.), *An integrated approach to communication theory and research*. Mahwah, NJ: Lawrence Erlbaum Associates, (pp.93-110).

McCombs, M. E., & Shaw, D. L. (1972). The agenda-setting function of mass media. Public Opinion Quarterly, 36(2), 176.

Naaman, M., Becker, H., & Gravano, L. (2011). Hip and trendy: Characterizing emerging trends on Twitter. *Journal of the American Society for Information Science and Technology, 62*(5), 902-918.

Naaman, M., Boase, J., & Lai, C.-H. (2010). Is it really about me?: Message content in social awareness streams. In K. I. Quinn, C. Gutwin & J. C. Tang (Eds.), *Proceedings of the 2010 ACM conference on Computer supported cooperative work*. New York, NY: ACM Press.

Nakasaki, H., Kawaba, M., Utsuro, T., & Fukuhara, T. (2009). Mining cross-lingual/cross-cultural differences in concerns and opinions in blogs. *Lecture Notes in Computer Science, 5459/2009*, 213-224.

Naveed, N., Gottron, T., Kunegis, J., & Alhadi, A. C. (2011). Bad news travel fast: A content-based analysis of interestingness on Twitter. *WebSci 2011*, Retrieved July 16, 2011 from: http://www.websci2011.org/fileadmin/websci/Papers/2050_paper.pdf.

Neuendorf, K. (2002). *The content analysis guidebook.* London: Sage.

Pak, A., & Paroubek, P. (2010). Twitter as a corpus for sentiment analysis and opinion mining. In *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC 2010)* (pp. Retrieved October 5, 2010 from: http://www.lrec-conf.org/proceedings/lrec2010/pdf/2385_Paper.pdf). Malta: European Language Resources Association.

Phelan, O., McCarthy, K., & Smyth, B. (2009). Using twitter to recommend real-time topical news. In *Proceedings of the third ACM conference on Recommender systems (RecSys '09)* (pp. 385-388). New York, NY: ACM Press.

Riegert, K. (2011). Pondering the future for foreign news on national television. *International Journal of Communication, 5*, Retrieved December 7, 2011 from: http://www.ijoc.org/ojs/index.php/ijoc/article/view/2488.

Romero, D. M., Meeder, B., & Kleinberg, J. (2011). Differences in the mechanics of information diffusion across topics: idioms, political hashtags, and complex contagion on Twitter. In *Proceedings of the 20th international conference on world wide web (WWW '11)* (pp. 695-704). New York, NY: ACM.

Sakaki, T., Okazaki, M., & Matsuo, Y. (2010). Earthquake shakes Twitter users: real-time event detection by social sensors. In *Proceedings of the 19th international conference on World wide web*. New York, NY: ACM Press.

Signorini, A., Segre, A. M., & Polgreen, P. M. (2011). The use of Twitter to track levels of disease activity and public concern in the U.S. during the influenza A H1N1 pandemic. *PLoS ONE, 6*(5), e19467.

Subašić, I., & Berendt, B. (2011). Peddling or creating? Investigating the role of Twitter in news reporting. *Lecture Notes in Computer Science, 6611/2011*, 207-213.

Takhteyev, Y., Gruzd, A., & Wellman, B. (2012). Geography of Twitter networks. *Social Networks*, 34(1), 73-81

Thelwall, M., Buckley, K., & Paltoglou, G. (2011). Sentiment in Twitter events. *Journal of the American Society for Information Science and Technology, 62*(2), 406-418.

Thelwall, M., & Prabowo, R. (2007). Identifying and characterising public science-related concerns from RSS feeds. *Journal of the American Society for Information Science & Technology, 58*(3), 379-390.

Thelwall, M., Wouters, P., & Fry, J. (2008). Information-centred research for large-scale analysis of new information sources. *Journal of the American Society for Information Science and Technology, 59*(9), 1523-1527.

Tumasjan, A., Sprenger, T. O., Sandner, P. G., & Welpe, I. M. (2010). Predicting elections with Twitter: What 140 characters reveal about political sentiment. In *Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media* (pp. 178-185). Menlo Park, CA: The AAAI Press.

Twitter (2011). Hot topics. Retrieved December 7, 2011 from: http://yearinreview.twitter.com/en/hottopics.html.

Wang, A. H. (2010). Don't follow me: Spam detection in Twitter. In *Proceedings of the International Conference on Security and Cryptography (SECRYPT)* (pp. Retrieved October 5, 2010 from: http://ceas.cc/2010/papers/Paper%2021.pdf).

Wanta, W., Golan, G. J., & Lee, C. (2004). Agenda setting and international news: Media influence on public perceptions of foreign nations. *Journalism and Mass Communication Quarterly*, 81(2), 364-377.

Westman, S., & Freund, L. (2010). Information interaction in 140 characters or less: Genres on twitter. In *Proceedings of the third symposium on Information interaction in context (IIiX '10)*. New York, NY: ACM Press.

Wigand, F. D. L. (2010). Twitter in government: building relationships one Tweet at a time. In *Proceedings of the Seventh International Conference on Information Technology* (pp. 563-567). Los Alamitos, CA: IEEE.

Winsten, J.A. (2000). The Harvard Alcohol Project: Promoting the "Designated Driver". In: M. Suman & G. Rossman, Advocacy groups and the entertainment industry (pp. 3-8). Westport, CT: Praeger Publishers.

Wu, H. D. (2000). Systemic determinants of international news coverage: a comparison of 38 countries. *Journal of Communication, 50*(2), 110-130.

Zaller, J. & Chiu, D. (1996). Government's little helper: US press coverage of foreign policy crises 1945–1991, *Political Communication*, 13(4), 385-405.

Zhang, M., Jansen, B. J., & Chowdhury, A. (2011). Influence of business engagement in online word-of-mouth communication on Twitter: A path analysis. *Electronic Markets: The International Journal on Networked Business*, 21(3), 161-175.

Zhao, D., & Rosson, M. B. (2009). How and why people Twitter: The role that micro-blogging plays in informal communication at work. In *Proceedings of the ACM 2009 international conference on Supporting group work (GROUP '09)* (pp. 243-252). New York: ACM Press.

Zhao, W. X., Jiang, J., Weng, J., He, J., Lim, E.-P., Yan, H., et al. (2011). Comparing Twitter and traditional media using topic models *Lecture Notes in Computer Science, 6611/2011*, 338-349.