

Final Draft of accepted paper: Tausczik, Y. R., Faase K., Pennebaker, J. W., & Petrie, K. J. (in press). Public anxiety and information seeking following H1N1 outbreak: Weblogs, newspaper articles and Wikipedia visits. Health Communication.

Public anxiety and information seeking following the H1N1 outbreak:

Web blogs, newspaper articles and Wikipedia visits

### Abstract

Web-based methodologies may provide a new and unique insight into public response to an infectious disease outbreak. This naturalistic study investigates the effectiveness of new web-based methodologies in assessing anxiety and information-seeking in response to the 2009 H1N1 outbreak by examining language use in web blogs, newspaper articles, and web-based information seeking. Language use in blogs and newspaper articles was assessed using Linguistic Inquiry and Word Count, and information-seeking was examined using the number of daily visits to H1N1-relevant Wikipedia articles. The results show that blogs mentioning 'swine flu' used significantly higher levels of anxiety, health, and death words and lower levels of positive emotion words than control blogs. Change in language use on blogs was strongly related to change in language use in newspaper coverage for the same day. Both the measure of anxiety in blogs mentioning 'swine flu', and the number of Wikipedia visits followed similar trajectories, peaking shortly after the announcement of H1N1 and then declining rapidly. Anxiety measured in blogs preceded information-seeking on Wikipedia. These results show that the public reaction to H1N1 was rapid and short-lived. This research suggests that analysis of web behavior can provide a source of naturalistic data on the level and changing pattern of public anxiety and information-seeking following the outbreak of a public health emergency.

Key words: H1N1, internet, health anxiety, information-seeking, language

Public anxiety and information seeking following the H1N1 outbreak:

Web blogs, newspaper articles and Wikipedia visits

Influenza A/H1N1 became world news following April 22, 2009 when the virus emerged in Mexico and the United States and spread rapidly into other countries. News of the virus, commonly known as 'swine flu', created considerable initial public anxiety as H1N1 was labelled a global pandemic by the World Health Organization (WHO). Countries quickly instituted measures aimed at limiting the spread of the virus, encouraging the public to increase hygiene precautions through hand washing and cleaning household surfaces, as well as staying away from work or public places after contracting flu-like symptoms.

As the health belief model predicts (Janz & Becker, 1984), the willingness of individuals to engage in containment and preventive behaviors such as hand hygiene and getting vaccinated against H1N1 is strongly associated with anxiety concerning the illness (Leung et al., 2005), perceptions about the severity of the illness and the perceived effectiveness of the recommended behaviors (Lau et al., 2009). Knowledge of how the public views disease outbreaks is therefore critical for understanding how likely individuals are to adopt recommended precautions and for informing the types of public health information messages that should be broadcast (Rubin, Amlôt, Rogers, Hall, Leach, Simpson & Wessely, 2010; Rubin et al., 2007). Within the agenda setting literature, researchers have begun to monitor web-based behavior, such as online discussions and aggregate searches, to track public reaction to news stories (Roberts, Wanta, & Dzwo, 2002; Weeks & Southwell, 2010). They have found that public reactions are visible faster in web activity than more traditional assessment methods.

The traditional approach to monitoring public perceptions of a health threat is to undertake a cross-sectional telephone survey (e.g. Leung et al., 2005; Lau et al., 2009; Lau,

Kim, Tsui, & Griffiths, 2007; Rubin, Amlôt, Page, & Wessely, 2009). Such surveys have provided valuable information about both the public perception of the H1N1 virus and the recommended actions to reduce the spread of the illness. In a United Kingdom telephone survey conducted just over two weeks after the H1N1 outbreak, Rubin and colleagues found anxiety about outbreak to be low, with only 2% reporting high anxiety and less than 5% reporting they had avoided public places due to the H1N1 outbreak (Rubin et al., 2009). The limitation of telephone surveys is they provide only a narrow window into public perceptions at the time the data is gathered and do not pick up changes over the course of a health emergency unless the survey is continuously updated, which is costly (Groves, 1990).

Recently, new web and text analysis applications have become available that offer the opportunity to assess public responses to health threats and other public health issues in real time by monitoring web activity (Freifeld, Mandl, Reis & Brownstein, 2008). As expected, the outbreak of H1N1 caused a large increase in activity on social networking websites and in blogs. For example, Twitter messages or 'tweets' containing the term 'swine flu' rose from almost zero to 125,000 per day by May 1, 2009. Web blogs showed a similar pattern with a large increase in the percentage of blogs mentioning 'swine flu' (Petrie & Faasse, 2009). Web applications like Trendrr.com, Blogpulse, google.com/trends, wikistats and spinn3r.com can monitor the frequency certain phrases and terms appear on the web or are searched for by internet users. Google search terms such as 'flu remedy' and 'antiviral medication' have been found to be closely associated with visits by patients to physicians with seasonal flu symptoms (Ginsberg et al., 2009). A recently developed text analysis program, Linguistic Inquiry and Word Count (LIWC), provides an analysis of written text by classifying words into various psychologically meaningful linguistic categories (Pennebaker, Booth, & Francis, 2007). This program has been used to track changes in emotional responses to disasters through analysis of blogs and on-line diaries (Cohn, Mehl, & Pennebaker, 2004). LIWC has been shown to be a valid measure of emotional expression in natural language. There is

high reliability between LIWC results and human ratings of online texts as well as congruence between LIWC results and texts written while subjects were primed to express different emotions (Bantum & Owen, 2009; Kahn, Tobin, Massey, & Anderson, 2007).

We were interested in whether web-based text analysis and new internet metrics offer an additional means of monitoring public anxiety following a public health emergency. This new approach compliments traditional measures that rely on self-report because anxiety is assessed indirectly through language use and online behavior. By assessing anxiety indirectly researchers can avoid self-report biases from social desirability effects or memory distortions inherent in retrospective accounts. It is also may be a cost effective research strategy when time or money is unavailable for a more thorough cross-sectional telephone study.

By using both the blog tracking application Spinn3r and the text analysis program LIWC we investigated in this study how public anxiety concerning H1N1 changed over the first two weeks following the outbreak. We also investigated how changes in the use of anxiety related words in blogs were related to swine flu-related information seeking behavior on Wikipedia. In particular, articles on influenza or flu, influenza symptoms, or influenza prevention.

### Method

We monitored public anxiety by examining the language used in personal blog entries published in the two weeks that followed the first WHO announcement of the presence of Swine Influenza A/H1N1 cases in Mexico and the United States on April 24, 2009. The period of two weeks was chosen because by the end of this period the number of blog entries discussing H1N1 had returned to a low baseline level. We compared personal blog entries that mentioned 'swine flu' to those that did not include these target words in order to assess differences between these two groups in the discussion of health concerns and emotions over time. Language use in blog entries was then compared with language

use in newspaper articles over the same time period. We also examined information-seeking behavior, as indicated by the number of visits to H1N1-related Wikipedia pages, in the two weeks prior and after the WHO announcement of the H1N1 outbreak. Wikipedia is in the top 10 most visited websites (Alexa, 2009), and is a popular health reference website (Laurent & Vickers, 2009).

### *Blog Data*

Blog entries were gathered between April 24, 2009 and May 7, 2009 using the Spinn3r database. Spinn3r is a blog aggregating service that monitors weblogs and stores the main content of a blog entry. Spinn3r has the ability to filter spam and limit searches to English-only blogs. To limit the sample to blog entries maintained by individuals, entries were selected only from the main personal blog domains: blogspot, wordpress and livejournal. Outside content selected by bloggers for inclusion in their blogs was retained in the data set, as these words were selected by the bloggers to convey a message to their audience. Blogs published directly by major news organizations were manually removed.

All blog entries that mentioned 'swine flu' were collected. The number of blog entries that mentioned swine flu varied by day, and an equal number of blog entries that did not mention these words were randomly drawn per day as a comparison group. Random blog selection was achieved by randomizing the order of minutes in each day, and blog entries published during the minutes on the list were drawn until the appropriate sample size was reached. A total of 9,508 blog entries that mentioned 'swine flu' were collected. The smallest number of 'swine flu' blogs per day was 102, and the maximum was 2,105 ( $M = 679.1$ ,  $SD = 543.5$ ). An equal number of control blog entries that did not mention 'swine flu' were collected each day, resulting in a total of 19,016 blog entries. The overall mean number of words per blog entry was 506.6 ( $SD = 1,254.5$ ).

### *Newspaper Articles*

English language newspaper articles that mentioned 'H1N1' or 'swine flu' were gathered between April 24, 2009 and May 7, 2009 from the LexisNexis database. A total of 4040 newspaper articles mentioning 'swine flu' were collected. The smallest number of articles per day was 9, and the maximum was 523 ( $M = 288.6$ ,  $SD = 180.3$ ).

#### *Wikipedia Article Visits Data*

Internet-based information seeking behavior data was drawn from 39 articles that were referenced through in text links on the English Wikipedia page for 'Swine Influenza.' Articles were selected based on their relevance to swine flu because they were either influenza-related (e.g. 'influenza', 'H1N1', 'infectious disease'), prevention-related (e.g. 'surgical masks', 'hand sanitizer', 'Tamiflu'), or symptom-related (e.g. 'fever', 'fatigue', 'chills').

The number of visits per day to selected English Wikipedia articles were collected for the period from April 10, 2009 (2 weeks prior to the H1N1 outbreak announcement) to May 7, 2009. The number of visits, or page counts, are recorded for each Wikipedia article and made available to the public. There is daily fluctuation in the number of visits to Wikipedia articles based on the day of the week and this was considered as a factor in the analyses.

#### *Analysis*

To compare H1N1-related activity across the three different mediums, total numbers of blog entries, newspaper articles, and visits to relevant Wikipedia articles were tabulated by day and cross-correlations were computed between the different sources.

The text of the blog entries was cleaned for html tags and non-characters. The blog entries published on the same day were aggregated into a single file, by blog type (swine flu and control). All 28 text files, covering 14 days for 2 blog types, were processed using LIWC. The same procedure was conducted to process text from newspaper articles.

Words in blogs were grouped into the following linguistic categories: disease words (e.g. swine flu, H1N1), health-related words (e.g. fever, aching, doctor), anxiety words (e.g.

worried, upset, uneasy), death-related words (e.g. fatal, dying, dead) and positive emotion words (e.g. happy, hopeful, relieved). For the analyses, percentage of words in each category was then computed. LIWC has been shown to be a reliable and valid method for measuring psychological constructs, particularly emotion (Tausczik & Pennebaker, 2010). Multiple linear regression was used to test whether blog type, time, and the interaction between the two predicted language used in the web blog entries. Correlations were then computed to compare language use in blog entries with language use in newspaper articles.

To assess the impact of swine flu on Wikipedia page views, a 2(before, after) x 7(day of the week) repeated measures ANOVA was conducted on the recorded log transformed daily number of visits to an article. The number of visits were log transformed to meet the normality assumption. The day of the week was included as a factor to control for fluctuation in the number of daily visits to Wikipedia. A 2(first week, second week) x 7(day of the week) repeated measures ANOVA was conducted to examine log transformed number of visits in the two weeks after the WHO announcement of swine flu.

### Results

The total attention given to H1N1 shows a rapid increase, peak and decline over the two week period in newspaper articles, blog entries, and Wikipedia visits (Figure 1). This pattern is consistent with a singular, short-lived increase in public anxiety. Increased attention to H1N1 happens most rapidly in Wikipedia page views, then in the blogs, and finally in newspapers. The duration of peak attention to H1N1 is shortest for the blog writers, followed by Wikipedia viewers, and is longest in newspapers.

---

Insert figure 1 about here

---

Cross-correlations between the newspaper articles, blog entries, and Wikipedia visits show that all three are strongly related (Table 1). The number of blog entries was most strongly related to the number of newspaper articles and Wikipedia visits on the same day. The number of Wikipedia visits was most strongly related to the number of newspaper articles the following day. In other words, public reaction is visible in online information seeking before it is visible in the amount of newspaper coverage.

---

Insert table 1 about here

---

Language use in blog entries was analyzed over time as an indirect measure of how the public was reacting to news of the outbreak. Language use was compared in control and swine flu blog entries. In comparison to control blogs, swine flu blog entries had significantly higher use of health-related words, death-related words, and anxiety words, and had significantly lower use of positive emotion words (Table 1). More interestingly, the discussion of swine flu changed over the two-week period. There was a significant interaction between blog type and language use by day for health-related words ( $F_{time \times \text{blog type}}(3,24) = 189.2, p < 0.001, R^2 = 0.96$ ), death-related words ( $F_{time \times \text{blog type}}(3,24) = 43.43, p < 0.001, R^2 = 0.84$ ), anxiety words ( $F_{time \times \text{blog type}}(3,24) = 91.1, p < 0.001, R^2 = 0.92$ ), and positive emotion words ( $F_{time \times \text{blog type}}(3,24) = 95.98, p < 0.001, R^2 = 0.92$ ). References to health-related words, death-related words and anxiety decreased significantly over the two weeks, and there was a significant increase in the expression of positive emotion words over the same time period. There was no significant change in language use in the control blogs over time.

---

Insert table 2 about here

---

Language use in blog entries was compared to language use in newspaper articles. There were strong correlations between language use for newspaper articles and blog entries published on the same day for health-related words ( $r = 0.93$ ,  $p < 0.001$ ), death-related words ( $r = 0.82$ ,  $p < 0.001$ ), anxiety words ( $r = 0.71$ ,  $p < 0.01$ ) and positive emotion words ( $r = 0.95$ ,  $p < 0.001$ ). There was only a marginally significant relationship for the number of mentions of H1N1 or swine flu ( $r = 0.46$ ,  $p < 0.10$ ). The correlations were highest comparing language use in newspaper articles and blog entries on the same day. These results show there is a dramatic decrease in the intensity of language used to discuss H1N1 newspaper articles and blog entries from the first day of the announcement and the language used in newspaper articles and blog entries are strongly related.

The number of visits to 'swine flu'-related Wikipedia articles was analyzed over time as a behavioral measure of information seeking. As a validity check, the number of visits to Wikipedia were compared before and after the announcement of H1N1. As expected, there was an increase in visits to Wikipedia articles in response to H1N1. There were more visits to 'swine flu'-related Wikipedia articles in the two weeks after news of the outbreak than in the two weeks prior ( $F(1,38) = 39.8$ ,  $p < 0.001$ ). This provides support for the using the number of Wikipedia visits as a behavioral measure of public health anxiety. The increase in visits to Wikipedia pages happened within days of news of the outbreak and returned to baseline within a few weeks. The rise in number of visits in response to the epidemic was greater the first week than the second week ( $F(1,38) = 17.1$ ,  $p < 0.001$ ). At its peak, the seventh day, there were 11.94 times as many visits per article on average.

As demonstrated in Figure 1, using Wikipedia traffic or the number of blog entries related to H1N1 as a measure of total attention paid to H1N1 produces very similar results. However, using the language in blog entries or traffic on Wikipedia as an indirect measure of the public's perception of the severity of H1N1 produces very different results. Linguistic indicators peak the first day after the announcement, whereas there is the most traffic to

Wikipedia the seventh day after the announcement. This difference highlights the fundamental difference in these two measurements. Traffic on Wikipedia is an aggregate measure of the amount of total information seeking, whereas language use in blogs is a measure of the average individual's reaction. Even though, these represent measures at two different levels of aggregation--population versus individual—there are some comparisons that can be made.

Language use in 'swine flu'-related blog entries and traffic on 'swine flu'-related Wikipedia articles can be compared to their baseline values to calculate the most extreme value and the average rate of decrease in attention to and salience of H1N1. Baseline values were calculated by taking the mean number of visits to the Wikipedia articles in the two weeks before the announcement of H1N1 and by taking the mean language use in the control blog entries. Wikipedia traffic at its highest increased by 1,094% from baseline. This represented a 140% increase from baseline per day, until the peak, and then a 172% decrease per day. In comparison, language use in blog entries did not change as fast: health related words increased by 711%, death words increased by 231%, anxiety words increased by 226%, and positive emotion words decreased by 60%. For all language measures the most extreme value was the first day, there was a slower rate of decrease: health words decreased by 41% per day, death words decreased by 11% per day, anxiety words decreased by 14% per day, and positive emotion words increased by 3% per day. Although language use in blog entries returned to baseline at a slower rate than Wikipedia views, these results show that both increased by a dramatic amount and decreased rapidly.

#### Discussion

This study demonstrates that monitoring web-related activity in terms of blogs and reference sites visited online provided a dynamic picture of the way the public responded to the announcement of the H1N1 virus pandemic. We have provided evidence that changes in public anxiety about the virus can be detected in their on-line writing and in searching on

Wikipedia. Compared to control blogs, entries that mentioned 'swine-flu' showed a greater use of health-related words, death-related words, and anxiety words, and had significantly lower use of positive emotion words. We also found a rapid increase in people seeking information on 'swine-flu' immediately following the announcement of the outbreak.

Beyond demonstrating a new method for tracking public reactions to a health emergency, this study also provides evidence of the possible advantages of this approach over traditional or panel surveys. The first random digit dialing phone survey on H1N1, conducted two weeks after the initial news of the outbreak, found public anxiety levels to be much lower than expected (Rubin et al., 2009). One conclusion, from this study, was that the efforts to inform the public about what they could do to reduce the risk of H1N1 had eased public anxiety in response to the epidemic. However, our results suggest that there may have initially been significant public anxiety about H1N1 but it had already declined by the time this telephone survey was conducted. The dramatic increase in people seeking information on 'swine-flu' occurred immediately after the announcement during the first week and by the second week it had returned to near baseline. This shows one advantage of using on-line data is that it gives a much more complete and dynamic picture of changes in public anxiety than a single or even longitudinal panel survey.

The findings of this study also provide a picture of how new and traditional media react to a public health emergency. Previous studies suggest that online reactions to traditional media occur quickly and disappear quickly (Roberts et al., 2002; Weeks & Southwell, 2010). We also found these immediate effects between newspaper articles and blog entries. Language use in blog entries was highly correlated with language use in newspaper articles published the same day. Furthermore, the number of newspaper articles appearing on H1N1 was highly correlated with the number of blog entries on H1N1 on the same day. However, we also found that the amount of people seeking information about

H1N1 on Wikipedia preceded the attention given to H1N1 in newspaper articles. This highlights that internet information searching may be a particularly useful early marker of public anxiety. This finding is consistent with other work showing uncertainty around health often prompts individuals to seek out information (Rosen & Knäuper, 2009).

Tracking emotional expression and real world behavior using online activity and linguistic analysis provides a cost-effective and complementary alternative to traditional surveys and has the advantage of not relying exclusively on self-report but indirectly assessing public reactions through analysis of language and behavior. Future work could track a broader range of online behavior, such as changes in online purchases, changes in topics on social bookmarking sites like Digg.com, and the dissemination of health information through social networks sites like Facebook.com.

### *Limitations*

Automated language analysis has been shown to be an effective tool in detecting psychologically meaningful events. However, the simplicity of the method means that subtleties such as the subject of statements or sarcasm cannot be distinguished. Changes in language use could be about the emotions of other individuals, such as 'my husband is worried about swine flu' or in sarcastic statements, 'Gee, well, now I'm sure to panic'. Despite this limitation, the fact that bloggers were discussing swine flu by using language as emotionally laden as death-related and anxiety words means that these were salient concepts, whether or not the language was used literally.

Another caveat is that individual blog entries were collected and pooled rather than tracking language use in blog entries for individuals over time. It was rare that a blogger would consistently write about swine flu every day; this approach enabled a collection of entries that were all specifically related to swine flu. This approach allows potential selection biases; the blog entries about swine flu that were captured may have been written by bloggers who always expressed high levels of anxiety and concern over health topics. In

general, tracking public reaction indirectly through online texts and web activity is also subject to selection biases as not everyone has access to the internet and only some users actively participate by posting content (Beck, Yan & Wang, 2009).

Future research should be directed at examining of the validity of web activity data in comparison to traditional community surveys. There is also a need to compare the relationship between traditional self-report measures of emotions and behavior to web activity from the same sample of participants. While this study just looked at the effect of the announcement of H1N1, there is also the possibility of tracking reactions to multiple events at once as the story develops over time.

#### Conclusion

Analysis of web behavior can provide a useful indicator of the level and changing pattern of public anxiety and information-seeking following the outbreak of a health emergency. Real world behavior can be measured in a non-invasive cost-effective way to detect public reactions, which may be difficult using traditional survey methodologies. The findings of this study underscore the immediacy and volatility of public responses to public health emergencies and how web activity can be a useful tool to monitor these reactions over time.

## References

- Alexa the Web Information Company Page. <http://www.alexa.com/topsites>. Updated 24 November 2009. Accessed 24 November 2009.
- Bantum, E. O., & Owen, J. E. (2009). Evaluating the validity of computerized content analysis programs for identification of emotional expression in cancer narratives. *Psychological Assessment, 21*, 79-88.
- Beck, K.H., Yan, A.F., & Wang, M.Q. (2009). A comparison of web-based and telephone surveys for assessing traffic safety concerns, beliefs, and behaviors. *Journal of Safety Research, 40*, 377-81.
- Cohn, M. A., Mehl, M. R., & Pennebaker, J. W. (2004). Linguistic markers of psychological change surrounding September 11, 2001. *Psychological Science, 15*, 687-693.
- Freifeld, C.C., Mandl, K.D., Reis, B.Y., & Brownstein, J.S. (2008). HealthMap: Global infectious disease monitoring through automated classification and visualization of internet media reports. *Journal of the American Medical Informatics Association, 15*, 150-7.
- Ginsberg, J., Mohebbi, M. H., Patel, R. S., Brammer, L., Smolinski, M. S., & Brilliant, L. (2009). Detecting influenza epidemics using search engine query data. *Nature, 457*, 1012-1014.
- Groves, R.M. (1990). Theories and methods of telephone surveys. *Annual Review of Sociology, 16*, 221-40.
- Janz, N.K., & Becker, M.H. (1984). The health belief model: a decade later. *Health Education and Behavior, 11*, 1-47
- Kahn, J. H., Tobin, R. M., Massey, A. E., & Anderson, J. A. (2007). Measuring emotional expression with Linguistic Inquiry and Word Count. *The American Journal of Psychology, 120*, 263-286.

- Lau, J. T. F., Kim, J. H., Tsui, H. Y., & Griffiths, S. (2007). Anticipated and current preventive behaviors in response to an anticipated human-to-human H5N1 epidemic in the Hong Kong Chinese general population. *BMC Infectious Diseases, 7*, 18-30.
- Lau, J. T. F., Yeung, N. C. Y., Choi, K. C., Cheng, M. Y. M., Tsui, H. Y., & Griffiths, S. (2009). Acceptability of A/H1N1 vaccination during pandemic phase of influenza A/H1N1 in Hong Kong: population based cross sectional survey. *British Medical Journal, 339*, 1010-1018.
- Laurent, M. R., Vickers, T. M. (2009). Seeking health information online: Does Wikipedia matter? *Journal of the American Medical Informatics Association, 16*, 471-479.
- Leung, G. M., Ho, L. M., Chan, S. K. K., Ho, S. Y., Bacon-Shone, J., Choy, R. Y. L., et al. (2005). Longitudinal assessment of community psychobehavioral responses during and after the 2003 outbreak of severe acute respiratory syndrome in Hong Kong. *Clinical Infectious Diseases, 40*, 1713-20.
- Pennebaker, J. W., Booth, R. J., & Francis, M. E. (2007). *Linguistic Inquiry and Word Count: LIWC* [Computer software]. Austin, TX: [www.liwc.net](http://www.liwc.net).
- Petrie, K. J., & Faasse, K. (2009, June 11) Monitoring public anxiety about flu [Web log message]. Retrieved from <http://blogs.bmj.com/bmj/2009/06/11>.
- Roberts, M., Wanta, W., & Dzwo, T. H. (2002). Agenda setting and issue salience online. *Communication Research, 29*, 452-465.
- Rosen, N. O., Knäuper, B. (2009). A little uncertainty goes along way: State and trait differences in uncertainty interact to increase information seeking but also increase worry. *Health Communication, 24*, 228-238.
- Rubin, G. J., Amlôt, R., Page, L., & Wessely, S. (2009). Public perceptions, anxiety, and behavior change in relation to the swine flu outbreak: cross sectional telephone survey. *British Medical Journal, 399*, 144-156.
- Rubin, G.J., Page, L., Morgan, O., Pinder, R.J., Riley, P., Hatch, S., Maguire, H., Catchpole,

M., Simpson, J., & Wessely, S. (2007). Public information needs after the poisoning of Alexander Litvinenko with polonium-210 in London: cross sectional telephone survey and qualitative analysis. *British Medical Journal*, *335*, doi:10.1136/bmj.39367.455243.BE.

Tausczik, Y. R. & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology*, *29*, 24-54.

Weeks, B. & Southwell, B. (2010). The symbiosis of news coverage and aggregate online search behavior: Obama, rumors, and presidential politics. *Mass Communication and Society*, *13*, 341-360.

Table 1: Cross-correlations between number of newspaper articles, blog entries, and Wikipedia visits related to H1N1 in the first three days after the announcement.

	<b>Blog entries Day 1</b>	<b>Wikipedia visits Day 1</b>
<b>Newspaper articles Day 1</b>	0.67**	0.63*
<b>Day 2</b>	0.65*	0.89***
<b>Day 3</b>	0.47	0.86***

	<b>Newspaper articles Day 1</b>	<b>Wikipedia visits Day 1</b>
<b>Blog entries Day 1</b>	0.67**	0.72**
<b>Day 2</b>	0.63*	0.67*
<b>Day 3</b>	0.34	0.61*

	<b>Newspaper articles Day 1</b>	<b>Blog entries Day 1</b>
<b>Wikipedia visits Day 1</b>	0.63*	0.72**
<b>Day 2</b>	0.22	0.32
<b>Day 3</b>	0.25	-0.31

\*p < 0.05, \*\* p < 0.01, \*\*\* p < 0.001

Table 1: Language use between blog entries that discuss swine flu or controls.

	<b>Blog Entry Type</b>	
	Swine Flu	Control
	<u>M</u> ( <i>SD</i> )	<u>M</u> ( <i>SD</i> )
<b>Health-related words</b>	2.43% (1.0%)	0.52% (0.06%)
<b>Death-related words</b>	0.44% (0.10%)	0.20% (0.039%)
<b>Anxiety words</b>	0.55% (0.14%)	0.27% (0.018%)
<b>Positive emotion words</b>	2.11% (0.41%)	3.11% (0.13%)

\*All t-tests were significant with  $p < 0.001$  with 26 degrees of freedom

## Figure Caption

*Figure 1.* Changes in standardized number of newspaper articles, blogs, and Wikipedia visits in the first two weeks after H1N1 announcement. The dependent variables are z-scores where each unit represents 180 newspaper articles, 543 blog entries, or 3,571 Wikipedia visits per article.

