

Social Transmission and Viral Culture

JONAH BERGER

KATHERINE L. MILKMAN*

*Jonah Berger is an assistant professor of Marketing (jberger@wharton.upenn.edu) and Katherine L. Milkman is an assistant professor of Operations and Information Management (kmilkman@wharton.upenn.edu) at the Wharton School, University of Pennsylvania, Jon M. Huntsman Hall, 3730 Walnut Street, Philadelphia, PA 19104. Michael Buckley, Jason Chen, Michael Durkheimer, Henning Krohnstad, Heidi Liu, Lauren McDevitt, Areeb Pirani, Jason Pollack, and Ronnie Wang all provided helpful research assistance. Hector Castro and Premal Vora created the webcrawler that made this project possible. Devin Pope and Bill Simpson provided helpful suggestions on our analysis strategy. Thanks to Max Bazerman, John Beshears, Jonathan Haidt, Chip Heath, Yoshi Kashima, Dacher Keltner, Mark Schaller and seminar participants at the Association for Consumer Research Conference, Max's non-lab, Yahoo! Research, Singapore Management University, the Society for Personality and Social Psychology conference, and the University of Chicago for helpful comments on a prior version of the manuscript. The Wharton Interactive Media Initiative helped fund this research.

ABSTRACT

People often share news, opinions, and information, and social transmission shapes both individual behavior and collective outcomes. But why are certain things more viral than others? An analysis of over 7,500 *New York Times* articles published over six months suggests that individual-level psychological processes (e.g., emotion) act as a selection mechanism on culture, shaping what becomes viral. Even controlling for external drivers of attention (e.g., the time an article spent on the *Times*' homepage), awe-inspiring articles are more likely to be among the newspaper's most e-mailed stories on a given day. Practically useful, surprising, positive, and affect-laden articles are also more likely to be viral. The magnitudes of these relationships are considerable. These results underscore the importance of considering how individual-level psychological processes shape collective outcomes such as the transmission and prominence of culture.

KEYWORDS: Social Transmission, Word-of-Mouth, Social Epidemics

Social transmission is an integral part of everyday life. We forward *New York Times* articles to our friends, pass along YouTube videos to our relatives, and send restaurant reviews to our neighbors. Hardly a day goes by when we do not share news, opinions, or information with those around us.

Such transmission affects both individual behavior and culture more broadly. Interpersonal communication influences the attitudes people hold and decisions they make (Asch 1956; Katz and Lazarsfeld 1955; Sherif 1937). For example, it affects what products consumers buy (Godes and Mayzlin 2009; Leskovec, Adamic, and Huberman 2007) and what drugs doctors prescribe (Iyengar, Van den Bulte, and Valente 2010). It also shapes the spread of ideas (Heath, Bell, and Sternberg 2001), diffusion of innovations (Rogers 1995) and formation and prominence of culture (Fast, Heath, and Wu 2009; Kashima 2008).

But while it is clear that social transmission is frequent, and has important consequences, less is known about why certain things are shared widely while others are not. Some news stories spread like wildfire while others languish. Some rumors propagate broadly while others stagnate. As Sperber (1996) notes, “to explain culture... is to explain why and how some ideas happen to be contagious” (p. 1). But what separates highly shared or contagious cultural items from those that are less so?

In this article, we identify characteristics of cultural items that are highly shared. Rather than focusing on special people who may be particularly influential in the diffusion process (Katz and Lazarsfeld 1955, Merton 1968) or social network structures that are conducive to information spread (Watts 2002), we focus on the fit between cultural items and human psychology. Specifically, we analyze over six months of data

from the *New York Times* website (including over 7,500 articles) to examine whether shared psychological processes (i.e., emotion) are linked to whether an article makes the *New York Times*' most emailed list. Before describing our data and analysis, we first review literature on drivers of diffusion and then outline our cultural selection approach.

DRIVERS OF DIFFUSION

Most research on the drivers of diffusion has emerged out of sociology and marketing science and focused on either (a) aggregate social structures or (b) the positions of certain individuals within those structures. Building on research in physics and mathematics, a great deal of recent work has examined how social networks shape social epidemics and the spread of information (Van den Bulte and Wuyts 2007; Watts 2004; also see Brown and Reingen 1987). Certain network structures such as small worlds, for example, are more conducive to diffusion (Watts 2002).

Another stream of research has focused on how certain special individuals, whether through their position in a social network or otherwise, may be particularly influential in the diffusion process. Whether they are described as opinion leaders, connectors, social hubs, influencers, or influentials, decades of work suggest that a small subset of people play an important role in driving everyone else's behavior (Goldenberg, Han, Lehmann, and Hong 2009; Katz and Lazarsfeld 1955; Merton 1968; Rogers 1995; Van den Bulte and Joshi 2007). For example, people with large numbers of social ties, for example, are thought to be particularly influential because they can quickly share information with many others (Goldenberg, Han, Lehmann, and Hong 2009). However,

recent simulation research has raised questions about the importance of such special people (Watts and Dodds 2007). Rather than being driven by a few influential individuals, this work suggests that epidemics depend on a critical mass of easily influenced individuals.

These macro explanations for diffusion, however, tend to ignore how individual-level processes influence what gets shared (though see Frenzen and Nakamoto 1993). Focusing on network structure says little about *why* people share things, and focusing on the influence of special people provides little insight into why certain cultural items become viral while others do not. Similarly, while the notion of “easily influenced individuals” is compelling (Watts and Dodds 2007), it says little about when and why certain individuals might be easy to influence. In their early work on word-of-mouth, Brown and Reingen (1987) note that, “an enhanced understanding of social influence processes in consumer behavior may simply be obtained by examining which products or services consumers are more likely to ‘talk about’” (p. 361), yet little empirical work has answered this call.

CULTURAL SELECTION

In considering what type of content may be highly shared, we focus on how psychological processes may help shape collective outcomes (Berger and Le Mëns 2009; Conway and Schaller 2007; Gureckis and Goldstone 2009). A small but growing literature suggests that the survival and propagation of cultural items depends on their fit with shared psychological processes (Heath, Bell, and Sternberg 2001; Kashima 2008;

Schaller and Crandall 2004). Consequently content's transmission may depend on its ability to tap shared emotions or fit with transmitters' motivations and cognitive constraints (Kashima 2000; Peters, Kashima, and Clark 2009; Rubin 1995; Wojnicki and Godes 2008). Despite theorizing about such cultural selection, however, there is little evidence of a link between psychological processes and actual transmission in the field.

We attempt to address that gap by examining whether certain psychological characteristics of content are related to their virality. One intuitive reason why certain content may become viral is because it contains practically useful information. Stories about websites that offer good bargains, or foods that are rich in antioxidants, contain material that might be useful to a receiver. They have clear utility in the economic sense because they provides people with information that improves their lives. Indeed, while no research we are aware of has examined this issue empirically, this notion is consistent with theorizing in the rumors and urban legends literature (Allport and Postman 1947; Heath, Bell, and Sternberg 2001; Rosnow and Fine 1976). Further, because useful information has social exchange value (Homans 1958), sharing it may encourage reciprocity (Fehr, Kirchsteiger, Riedl 1998) or be driven by people's desire to look good or self-enhance (Wojnicki and Godes 2008).

Even beyond practical utility, however, we suggest that emotional aspects of content may drive transmission. People report discussing many of their emotional experiences with others, and the social sharing of emotion serves a variety of potential functions. First, emotional stimuli often elicit ambiguous sensations, and through talking about and sharing emotional content with others, people can clarify that ambiguity and gain deeper understanding of how they feel (Rime, Mesquita, Philippot, and Boca 1991).

Second, to the extent that emotional material challenges people's beliefs or way of seeing the world, they may share it with others to help them cope or reduce feelings of dissonance (Festinger, Riecken, and Schachter 1956). Third, sharing emotional content with others can strengthen social bonds and deepen social connections (Peters and Kashima 2007).

One emotion we focus on in particular is awe. Stimuli that open the mind to vast and often unconsidered possibilities can inspire awe, a unique human emotion that expands a reader's frame of reference (Keltner and Haidt 2003). Awe is the emotion of self-transcendence, a feeling of admiration and elevation in the face of something greater than the self (Haidt 2006). It occurs when two conditions are met (Keltner and Haidt 2003). First, people experience something vast: either physically vast such as the grand canyon, conceptually vast such as a grand theory or finding, or socially vast such as fame or power. Second, the vast experience cannot be accommodated by existing mental structures. Intellectual epiphanies, natural wonders, and great works of art can all make people feel a sense of awe (Shiota, Keltner, and Mossman 2007). Similarly, news stories about a treatment that may cure AIDS or a hockey goalie who continues to play even with brain cancer may both inspire some level of awe.

Awe may be linked to social transmission for a number of reasons. First, awe encourages people to connect with others and spread the word. People who have had epiphanies through drug use or religious experiences, for example, seem to have a deep need to talk about them or proselytize (James 1902; Keltner and Haidt 2003). Other types of awe-inducing experiences may activate the same psychological mechanisms evoked by epiphanies. Second, awe inducing stimuli also tend to be entertaining,

inspiring, and contain a great deal of information. Each one of these aspect in itself is a reason that people may share things. Third, because awe-inducing experiences are characterized by the accommodation of existing mental structures, they should be particularly likely to drive people to talk to others to understand how they feel (Rime, Mesquita, Philippot, and Boca 1991). Finally, awe-inducing experiences encourage people to look beyond themselves and deepen connections to the broader social world (Shiota, Keltner, and Mossman 2007). All of these factors suggest that awe should lead people to want to share.

Importantly, we distinguish awe from a variety of related dimensions that may also be linked to transmission. One such dimension is surprise. Surprise differs from awe in that it lacks vastness as an elicitor. While stories about someone who fell asleep in a bar overnight or Japanese farmers growing square watermelons to fit into refrigerators are surprising, the eliciting stimulus is not vast enough to generate awe. Due to its high entertainment value, however, surprising content may be viral. Sharing surprising information may also reflect positively on the sender, as it suggests that he or she knows interesting and unusual things. Indeed, though no work has looked at actual transmission, people report a greater willingness to share more surprising urban legends or social anecdotes (Heath, Bell, and Sternberg 2001; Peters, Kashima, and Clark 2009).

We also distinguish awe from valence more generally. People may share positive or upbeat stories, for example, because they make others feel better or boost their mood. While researchers have theorized about whether people tend to share positive or negative news, little empirical work has actually examined this question in the field. Indeed, researchers have noted that “more rigorous research into the relative probabilities of

transmission of positive and negative information would be valuable to both academics and managers,” (Godes et al. 2005, p. 419). Further, though some have suggested that there is more positive than negative word-of-mouth about brands (Keller and Libai 2009), even if this is the case it is unclear whether this is because of preferential transmission (i.e., people prefer to pass along positive information) or the distribution of events (i.e., there are more positive things that happen than negative ones). Thus without controlling for the distribution of information in a given context, it is impossible to know whether people are actually more likely to share positive news.

Overall then, the preceding discussion leads to the following hypotheses about psychological characteristics of content and virality. Given that no prior work has examined the link between awe and transmission, this is the main focus of our analyses.

H1: More awe-inspiring content will be more viral

Further, though not our primary focus, little work has actually empirically examined the other dimensions discussed. Consequently, by including them in our analyses, we not only test whether any relationship between awe and virality persists even controlling for these factors, but also provide insight into whether these factors themselves may be related to virality.

H2: More practically useful content will be more viral

H3: More surprising content will be more viral

H4: Positively valenced content will be more viral than negatively valenced content

H5: More affect laden content will be more viral

ANALYSIS OF THE *NEW YORK TIMES*' MOST EMAILED LIST

We investigate whether shared psychological processes shape what content is highly shared spreads by examining the virality of over 7,500 articles from one of the world's most popular newspapers: *The New York Times*. Because the *Times* covers a wide range of topics (from world news and politics to science and travel) and boasts the most frequented website of any newspaper in the United States (Nielson NetRatings NetView, 2008), it is an ideal venue for examining what types of content are frequently shared. The *Times* continually reports which 25 of its articles have been emailed most frequently in the last 24 hours, and we examine whether there is a link between psychological characteristics of articles (e.g., how awe inspiring or affect laden they are) and whether they make the *Times*' most emailed list.

In addition to quantifying the extent to which articles in our dataset inspire awe, provide practical utility, evoke surprise, contain emotion, and exhibit positive valence, we include a number of important control variables (separate from the content itself) in our analyses. Articles that appear on the front page of the physical paper or spend more time in prominent positions on the homepage, for example, may receive more attention and thus mechanically have a better chance of making the most emailed list. Consequently we control for these, and other potential external drivers of attention, to ensure that any relationship between content characteristics and virality are not the result of editorial decisions about what to feature or author fame.

Data

We collected information about every article written for the *Times* that appeared on the paper's homepage (www.nytimes.com) between August 30th 2008 and February 15th 2009 (7,710 articles). This data was captured by a webcrawler that visited the *Times* homepage every 15 minutes during the six months studied. The webcrawler recorded information about every article on the homepage and each article on the most emailed list (updated every 15 minutes). The content of AP, Reuters, and Bloomberg articles, as well as blogs, are not stored by the *Times*, and so was not available for our analyses. Videos and images with no text were also not included. We captured each article's title, full text, author(s), topic area (e.g., opinion or sports), and one sentence summary created by the *Times*. We also captured each article's section, page, and publication date if it appeared in the print paper, as well as the dates, times, locations and durations of all appearances it made on the homepage. 20% of the articles in our final data set eventually earned a position on the most e-mailed list.

Article Coding

We relied on human coders to quantify the extent to which each article contained practical information, inspired awe, or evoked surprise. Coders were blind to our hypotheses. They received the title and summary of each article, a web link to the article's full text, and detailed coding instructions (see Appendix). Given the overwhelming number of articles in our data set, we selected a random subsample for

coding (N = 3,000). Articles that earned a position on the most e-mailed list were oversampled to comprise 27% of the coded subsample. For each dimension, three independent raters rated each article (*Practical utility*: 1 = not at all practically useful, 5 = extremely practically useful; *Awe-inspiring*: 1 = not at all awe-inspiring, 5 = extremely awe-inspiring; *Surprise*: 1 = not at all surprising, 5 = extremely surprising). Raters were given feedback on their coding of a test set of articles until it was clear they understood the relevant construct. Inter-rater reliability was high on all dimensions ($\alpha_{awe} = .75$; $\alpha_{practicality} = .87$; $r_{surprise} = .77$), and scores were averaged across coders (see Table 1 for summary statistics) and then standardized. All uncoded articles were assigned a score of zero on each dimension after standardization, and a dummy was included in regression analyses to control for uncoded stories. This strategy allowed us to use the full set of articles collected to analyze the relationship between other content characteristics (that did not require manual coding) and virality.

Examples of highly awe-inspiring articles included: “Fury of Girl’s Fists Lifts Up North Korean Refugee Family” and “The Promise and Power of RNA”. Examples of highly practically useful articles included: “Voter Resources” and “It Comes in Beige or Black, but You Make It Green” (about being environmentally friendly when disposing of old computers). Finally, examples of highly surprising articles included: “Passion for Food Adjusts to Fit Passion for Running” (about a restaurateur who runs marathons) and “Pecking, but No Order, on Streets of East Harlem” (about chickens in Harlem)

Automated sentiment analysis was used to quantify the positivity (i.e., valence) and emotionality of each article. These methods are well-established (Pang and Lee 2008) and increase coding ease and objectivity. A computer program counted the

number of positive and negative words in each article using a list of 7,630 words classified as positive or negative by human readers (Pennebaker, Booth, and Francis 2007). Positivity was quantified as the difference between the percentage of positive and negative words in an article. Emotionality was quantified as the percentage of words that were classified as either positive or negative. These variables were also standardized to ease interpretation of our regression results (see Table 2 for correlations between standardized variables).

Examples of highly positive articles included: “Wide-Eyed New Arrivals Falling in Love With the City” and “Who Will Win the Super Bowl?”, while highly negative articles included: “Web Rumors Tied to Korean Actress’s Suicide” and “Germany: Baby Polar Bear’s Feeder Dies”. Examples of highly affect laden articles included: “Redefining Depression as Mere Sadness” and “When All Else Fails, Blaming the Patient Often Comes Next”.

Additional Controls

External factors (separate from content characteristics) may affect an article’s virality by functioning like advertising. Appearing earlier or in certain sections of the physical paper, spending more time in a prominent position on the homepage, being released when readership is greater, and being written by a famous author all likely generate attention for an article and increase its chances of making the most emailed list.

Consequently, we control for these factors. To characterize where an article appeared in the physical paper, we created dummy variables to control for the section

(e.g., A), and an indicator variable to control for the page (e.g., 1), where an article appeared in print. We also interact each section dummy with our page indicator variable to control for the possibility that appearing earlier in some sections has a different effect than appearing earlier in others. To characterize how much time an article spent in prominent positions on the homepage, we created variables that indicated where, when, and for how long every article was featured on the *Times* homepage. The homepage layout remained the same throughout the period of data collection. Articles could appear in several dozen positions on the homepage, so we aggregated positions into seven general regions based on locations that likely receive similar amounts of attention (see Figure 1). Variables indicating the amount of time an article spent in each region were included as controls after winsorization of the top 1% of outliers (to prevent extreme outliers from exerting undue influence on our results; see Tables A2 and A3 in the Appendix for summary statistics). We also created controls for the day of the week and the time of day (6 am – 6 pm or 6 pm – 6 am EST) when an article first appeared online.

We also control for author fame to ensure that our results are not driven by the tastes of particularly popular writers whose stories may be particularly likely to be shared. To quantify author fame, we capture the number of Google hits returned by a search for the article's first author's full name (as of February 15, 2009). Due to its skew, we use the logarithm of this variable as a control in our analyses.

We also control for writing style (e.g., complexity and author gender) and article length, as both might influence transmission and affect the likelihood that an article contains awe-inspiring content. We control for writing complexity, or how difficult a piece of writing is to read, using the *SMOG Complexity Index* (McLaughlin 1969). Since

male and female authors have different writing styles (Koppel, Argamon, and Shimoni 2002; Milkman, Carmona and Gleason, 2007), we control for the gender of an article's first author (male, female or unknown due to a missing byline). We classify gender using a first name-gender mapping list from prior research (Morton and Zettelmeyer 2003). For names that were classified as gender neutral or did not appear on this list, research assistants determined author gender by looking up the authors online. Finally, we control for article length in words. Longer articles may be more likely to go into enough detail to inspire awe but may simply be more viral because they contain more information.

Analysis Strategy

To analyze the relationship between an article's content characteristics and the likelihood that it will make the *New York Times*' most e-mailed list, we use the following logistic regression specification:

$$(1) \quad \text{makes_it}_a = \frac{1}{1 + \exp \left\{ - \left[\begin{array}{l} \alpha + \beta_1 * z\text{-awe-inspiring}_a + \beta_2 * z\text{-surprise_score}_a + \\ \beta_3 * z\text{-positivity}_a + \beta_4 * z\text{-emotionality}_a + \\ \beta_5 * z\text{-practicality}_a + \theta * X_a \end{array} \right] \right\}}$$

where makes_it_a is a variable that takes on a value of one when an article a earns a position on the most e-mailed list and zero otherwise. Our primary predictor variables quantify the extent to which an article a was coded as practically useful, surprising, positive, emotional or awe-inspiring. X_a is a vector of the other control variables described above¹,

¹ This includes including: indicators of the number of hours an article spent in each of seven online locations, dummies indicating the day of the week when the article first appeared online (Monday omitted),

RESULTS

Results indicate that content's ability to inspire awe is strongly associated with its virality. Supporting Hypothesis 1, Model 1 illustrates that stories that open the mind and inspire awe are more likely to make the *New York Times*' most emailed list (Table 3). This is the case even controlling for an article's practical value as well as how surprising it is, its valence, and its degree of emotionality more generally (Table 3, Model 2). Including these controls distinguishes the effect of awe from other related constructs. Importantly, however, these factors are also linked to virality. That is, supporting Hypotheses 2-5, more surprising, practically useful, emotion-laden, or positive articles are more likely to make the most emailed list. Ancillary analyses also show that the strong positive relationship between awe and virality persists even controlling for the amount of interest an article evokes or alternate measures of awe (see ancillary analyses below).

The relationship between awe, as well as the other psychological characteristics investigated, and virality is also robust to controlling for how prominently an article is featured. As predicted, both where and when an article is featured are linked to virality, but even controlling for these factors, awe-inspiring articles are still more likely to make the most emailed list (as are surprising, practically useful, emotion laden, and positive

a dummy indicating whether the article first appeared online at night (6 pm – 6 am EST), a dummy indicating which section in the physical paper the article appeared in, an indicator of the page number in that section in the physical paper that the article appeared in, interactions between physical paper section dummies and the page number indicator, the first author's fame score, the article's standardized complexity score, dummies indicating whether the first author is female or of unknown gender (due to a missing byline), the article's wordcount, and a dummy indicating whether the article in question was one of the 3,000 coded manually on the characteristics: *practicality*, *surprise*, and *awe-inspiring*.

articles: Table 3, Model 3). The robustness of our results to the inclusion of controls for “advertising” in the *Times* ensures that the high observed transmission rates of awe-inspiring stories, for example, are not due to an editorial preference for prominently featuring awe-inspiring news, which could mechanically increase the virality of such items. Longer articles, articles by more famous authors and articles written by women are also more likely to make the most emailed list, but controlling for these factors does not meaningfully change the relationship between awe and virality (or the relationships between other psychological characteristics of content and virality: Table 3, Model 3).

Further, though external drivers of attention (e.g., being prominently featured) are important to cultural success, our results indicate that the fit between cultural items and shared human psychology is of similar importance. Parameter estimates imply that a one standard deviation increase in an article’s awe-inspiring score, for example, increases the odds that it will make the most e-mailed list by a factor of 1.5 (Figure 2). This increase is equivalent to the effect of spending an additional 4.3 hours as the lead story on the *Times* website, which is over six times the average number of hours the average article spends in that position.

Our results are also robust to controlling for an article’s general topic (i.e., 20 areas classified by the *Times* such as opinion or health: Table 3, Model 4). This indicates that our findings are not merely driven by certain topic areas (e.g., science or health) that may tend to score highly on certain psychological dimensions (i.e., surprising or awe-inspiring) and be particularly likely to make the most e-mailed list. Rather, this more conservative test of our hypothesis demonstrates that the observed effects of psychological characteristics hold not only across topics but also within them. Even

among opinion or health articles, for example, awe-inspiring articles or surprising articles are more likely to be viral.

Ancillary Analyses

Our results are also robust to different ways of quantifying the extent to which an article inspires awe. One such measure relied on automated textual analysis (instead of manual coding). Manual coding is useful because it can pick up on subtle characteristics of an article, but one could argue that this measure is overly subjective and that a more objective measure (like a count of the number of words related to awe in each article) might be superior. While we are not aware of a word list that exactly matches this criterion, the General Inquirer's textual analysis software (<http://www.wjh.harvard.edu/~inquirer/>) contains a category of words related to enlightenment. When the percentage of such words per article is used as a proxy for awe, we find results consistent with those reported in Table 1. This alternate measure of awe is significantly associated with an article's likelihood of making the *New York Times*' most emailed list, and this relationship persists even when all of our main controls are added to the analysis. It is important to note that while this word count measure may seem more objective than the measure we rely on in our primary analyses, it may be subject to other flaws. That said, the fact that our results are similar even when using this alternate measure, which takes into account the full text of the article, underscores the link between awe and virality. Our results are also robust using an alternate hand-coded measure of awe (see Appendix), though the magnitude of the estimated effects is slightly

reduced. Finally, our results are robust to including a hand-coded control for how interesting an article is.

The results presented in Table 3 also remain meaningfully unchanged in terms of magnitude and statistical significance if we: (1) add squared and/or cubed terms quantifying how long an article spent in each of seven homepage regions; (2) add dummies indicating whether an article ever appeared in a given homepage region; (3) split the region variables into time spent in each region during the day (6 am – 6 pm EST) and night (6 am – 6 pm EST); (4) control for the day of the week when an article was published in the physical paper (instead of online); (5) winsorize the top and bottom 1% of outliers for each control variable in our regression; (6) or include day of the year fixed effects and clustering standard errors by day of the year. These checks indicate that the observed results are not an artifact of the particular regression specifications we rely on in our primary analyses.

GENERAL DISCUSSION

A broad analysis of social transmission suggests that psychological processes help shape what becomes viral. More awe-inspiring articles are more likely to make the *New York Times*' most emailed list, as are articles that are more surprising, positively valenced, emotion-laden, or practically useful. By demonstrating the strength of these relationships outside the laboratory, and across a diverse body of content, we provide evidence that psychological characteristics of content play an important role in shaping virality in the field. This naturalistic setting also allows us to test the importance of

psychological factors relative to other potential drivers of virality (e.g., being featured prominently). Our findings underscore the notion that individual-level psychological processes play an important role in shaping collective outcomes.

These findings also help shed light on the ongoing debate about whether people tend to share positive or negative news. Common wisdom seems to suggest that people are more likely to pass along negative news. Further, though some have suggested that positive word-of-mouth is more prevalent (Keller and Libai 2009), this does not necessarily mean that people are more likely to share positive news. Finally, while some work has investigated how the valence of news people prefer to share may vary by domain (Heath 1996) or expertise (Wojnicki and Godes 2008), we are not aware of any investigation that has examined the virality of actual news across a wide variety of domains in the field. Our results suggest that in general, positive news is more viral. This does not mean people never transmit negative content, and in ancillary analyses we find some suggestion that specific negative emotions may be linked to transmission (i.e., anxiety, also see Heath, Bell and Sternberg 2001 for the link between disgust and sharing urban legends). In general, however, our findings suggest that people prefer to share positive rather than negative news.

More broadly, our results suggest that transmission is not just about value exchange, but also about deepening social connections. Consistent with social exchange theory (Homans 1958) and the notion that people may share word-of-mouth to self enhance (Wojnicki and Godes 2008), useful articles are more likely to be viral. Surprising content is also more viral, a pattern that may also be driven by self enhancement motives, as sharing surprising content allows people to signal that they

know unexpected things. Even after controlling for these effects, however, awe-inspiring, and emotional content are both more likely to make the most emailed list. While such content does not clearly produce immediate economic value in the traditional sense, sharing affectively rich content can reinforce shared views and deepen social bonds (Heath, Bell and Sternberg 2001; Peters and Kashima 2007). Thus people may not only share for self-enhancement purposes or to generate reciprocity, but also to deepen connections with others.

In considering how psychological processes shape what becomes viral, our findings dovetail with recent theorizing on the psychological foundations of culture (Kashima 2008; Schaller and Crandall 2004). Research on cross-cultural psychology has examined how cultural background (e.g., American vs. East Asian) shapes psychological process and individual behavior (Markus and Kitayama 1991; Nisbett, Peng, Choi, and Norenzayan 2001), but the influence is reciprocal. Individual behavior, and the psychological processes that underlie it, shape the beliefs, norms, symbols, and institutions that make up culture (Heath, Bell, and Sternberg 2001; Schaller Conway, and Tanchuk 2002). Supporting this premise, our work shows a link between psychological characteristics and whether content become highly shared.

Future research might examine how the effects observed here are moderated by situational or relationship factors. Given that the weather can affect people's moods, for example, it may affect the type of content that gets shared. People might be more likely to share positive stories on overcast days, for example, to make others feel happier. Alternatively, people might be more likely to share more negative stories on overcast days due to mood congruence. More broadly, other cues in the environment might

change what people share by making certain topics more accessible (Berger and Fitzsimons 2008; Nedungadi 1990). If the Yankees win the World Series, for example, that should make front page news, but as a result, people may also be likely to share any sports story more generally because that topic is top of mind. Moderators might also occur at the level of the transmission dyad. People share different things to different people, and what they share depends on the recipient and relationship characteristics (Stephen and Lehmann 2009). Consequently, it would be interesting to examine how the type of content interacts with the relationship characteristics to determine what is shared.

In conclusion, truly understanding collective outcomes requires integrating disciplinary perspectives. As it stands, separate research traditions examine social epidemics at the aggregate level and transmission at the individual level. Gaining richer insight into these complex, multiply-determined phenomena requires not only examining individual mechanisms (i.e., social networks or psychological drivers of transmission) in isolation, but also understanding how they combine to drive collective outcomes.

REFERENCES

- Allport, Gordon W. and Leo Postman (1947), *The Psychology of Rumor*, New York: Henry Holt.
- Asch, Solomon E. (1956), "Studies of independence and conformity: A minority of one against a unanimous majority," *Psychological Monographs*, 70 (416).
- Bentley, R. Alexander and Stephen J. Shennan (2005), "Random copying and cultural evolution," *Science*, 309, 877-879.
- Berger, Jonah and Gael Le Mens (2009), "How Adoption Speed Affects the Abandonment of Cultural Tastes," *Proceedings of the National Academy of Sciences*, 106, 8146-8150.
- Berger, Jonah and Gráinne M. Fitzsimons (2008), "Dogs on the Street, Pumas on Your Feet: How Cues in the Environment Influence Product Evaluation and Choice," *Journal of Marketing Research*, 45(1), 1-14.
- Brown, Jacqueline Johnson and Peter H. Reingen (1987), "Social Ties and Word-of-Mouth Referral Behavior," *Journal of Consumer Research*, 14, 350-62.
- Conway, Lucian Gideon, III, and Mark Schaller (2007), "How communication shapes culture," in *Social Communication*, ed. K. Fiedler, New York: Psychology Press, 107-127.
- Fast, Nathaneal J., Chip Heath, and George Wu (2009), "Common ground and cultural prominence: How conversation reinforces culture," *Psychological Science*, 20, 904-911.

- Godes, David, Dina Mayzlin, Yubo Chen, Sanjiv Das, Chrysanthos Dellarocas, Bruce Pfeiffer, Barak Libai, Subrata Sen, Mengze Shi and Peeter Verlegh (2005), "The Firm's Management of Social Interactions," *Marketing Letters*, 16 (3/4), 415-428.
- Godes, David and Dina Mayzlin (2009), "Firm-Created Word-of-Mouth Communication: Evidence from a Field Study," *Marketing Science*, 28(4), 721-39.
- Fehr, Ernst, Georg Kirchsteiger and Arno Riedl (1998), "Gift Exchange and Reciprocity in Competitive Experimental Markets," *European Economic Review*, 42, 1, 1-34.
- Festinger, Leon, Henry W. Riecken, and Stanley Schachter (1956), *When Prophecy Fails*, New York: Harper and Row.
- Frenzen, Jonathan and Kent Nakamoto (1993), "Structure, Cooperation, and the Flow of Market Information", *Journal of Consumer Research*, 20, 360-375
- Goldenberg, Jacob, Sangman Han, Donald R. Lehmann, and Jaewon Hong (2009), "The role of hubs in the adoption process," *Journal of Marketing*, 73 (March), 1-13.
- Goodman, John (1999), "Basic Facts on Customer Complaint Behavior and the Impact of Service on the Bottom Line," *Competitive Advantage*, June 1999, 1-5.
- Gureckis, Todd M. and Robert L. Goldstone (2009), "How to name your child: Individual Decision-making and the Social Construction of Value," *TopiCS in Cognitive Science*, 1, 456-478.
- Haidt, Jonathan (2006), *The happiness hypothesis: Finding modern truth in ancient wisdom*, New York: Basic Books.
- Heath, Chip (1996), "Do people prefer to pass along good news or bad news? Valence and relevance of news as a predictor of transmission propensity," *Organizational Behavior and Human Decision Process*, 68 (2) (November), 79-94.

- Heath, Chip, Chris Bell, and Emily Sternberg (2001), "Emotional selection in memes: the case of urban legends," *Journal of Personality and Social Psychology*, 81, 1028-1041.
- Homans, George C. (1958), "Social behavior as exchange," *American Journal of Sociology*, 63 (6), 597-606.
- Iyengar, Raghuram, Christophe Van den Bulte, and Thomas W. Valente (2008), "Opinion Leadership and Social Contagion in New Product Diffusion," *Marketing Science Institute Research Report*, 120 (No. 8).
- James, William (1902). *The Varieties of Religious Experience*.
- Kashima, Yoshihisa (2000), "Maintaining cultural stereotypes in the serial reproduction of narratives," *Personality and Social Psychology Bulletin*, 26, 594-604.
- Kashima, Yoshihisa (2008), "A social psychology of cultural dynamics: examining how cultures are formed, maintained, and transformed," *Social and Personality Psychology Compass*, 2, 107-120.
- Katz, Elihu, Paul Felix Lazarsfeld, and Elmo Roper (1955), *Personal influence: The part played by people in the flow of mass communication*, Glencoe, IL: Free Press.
- Keller, Ed and Barak Libai (2009), "A Holistic Approach to the Measurement of WOM," *ESOMAR Worldwide Media Measurement Conference*, Stockholm.
- Keltner, Dacher and Jon Haidt (2003), "Approaching awe: a moral, spiritual, and aesthetic emotion," *Cognition and Emotion*, 17, 297-314.
- Koppel, Moshe, Shlomo Argamon, and Anat Rachel Shimoni (2002), "Automatically categorizing written texts by author gender," *Literary and Linguistic Computing*, 17, 401-412.

- Leskovec, Jure, Lada A. Adamic, and Bernardo A. Huberman (2007), "The dynamics of viral marketing," *ACM Transactions on the Web*, 1 (1).
- Markus, Hazel Rose and Shinobu Kitayama (1991), "Culture and the Self: Implications for Cognition, Emotion, and Motivation," *Psychological Review*, 98 (2), 224-253.
- McLaughlin, G. Harry (1969), "SMOG grading: a new readability formula," *Journal of Reading*, 12, 639-646
- Merton, Robert K. (1968), "Patterns of Influence: Local and Cosmopolitan Influentials," in *Social Theory and Social Structure*, ed. Robert K. Merton, New York: Free Press, 441-74.
- Milkman, Katherine L., Rene Carmona, and William Gleason (2007), "A Statistical Analysis of Editorial Influence and Author-Character Similarities in 1990s New Yorker fiction," *Journal of Literary and Linguistic Computing*, 22, 305-328.
- Morton, Fiona Scott, Florian Zettelmeyer, and Jorge Silva-Risso (2003), "Consumer information and discrimination: does the Internet affect the pricing of new cars to women and minorities?" *Quantitative Marketing and Economics*, 1, 65-92
- Nedungadi, Prakash (1990), "Recall and Consumer Consideration Sets: Influencing Choice Without Altering Brand Evaluations," *Journal of Consumer Research*, 17 (December), 263-76.
- Nielson (2008). *NetRatings NetView*.
- Nisbett, Richard E., Kaiping Peng, Incheol Choi, and Ara Norenzayan (2001), "Culture and Systems of Thought: Holistic Versus Analytic Cognition," *Psychological Review*, 108 (April), 291-310.

- Pang, Bo and Lillian Lee (2008), "Opinion mining and sentiment analysis," *Foundations and Trends in Information Retrieval*, 2, 1-135.
- Pennebaker, James W., Roger J. Booth, and Martha E. Francis (2007), *LIWC2007: Linguistic Inquiry and Word Count*, Austin, Texas: liwc.net
- Peters, Kim and Yoshihasa Kashima (2007), "From social talk to social action: shaping the social triad with emotion sharing," *Journal of Personality and Social Psychology*, 93, 780-797.
- Peters, Kim, Yoshihasa Kashima, and Anna Clark (2009), "Talking about others: emotionality and the dissemination of social information," *European Journal of Social Psychology*, 39, 207-222
- Rime, Bernard, Batja Mesquita, Pierre Philippot, and Stefano Boca (1991), "Beyond the emotional event: Six studies on the social sharing of emotion," *Cognition and Emotion*, 5 (September-November), 435-465.
- Rogers, Everett M. (1995), *Diffusion of Innovations*, New York: Free Press.
- Rosnow, Ralph L. and Gary A. Fine (1976), *Rumor and gossip: The social psychology of hearsay*, New York: Elsevier.
- Rubin, David C (1995), *Memory in oral traditions: The cognitive psychology of epic, ballads, and counting-out rhymes*, Oxford, England: Oxford University Press.
- Salganik, Matthew J., Peter Sheridan Dodds, and Duncan J. Watts (2006), "Experimental study of inequality and unpredictability in an artificial cultural market," *Science*, 311, 854-856
- Schaller, Mark, Lucian Gideon Conway III, and Tracy L. Tanchuk (2002), "Selective pressures on the once and future contents of ethnic stereotypes: Effects of the

- communicability of traits,” *Journal of Personality and Social Psychology*, 82 (6) (June), 861-877.
- Schaller, Mark and Christian S. Crandall (2004), *The Psychological Foundations of Culture*, Mahwah, NJ: Lawrence Erlbaum Associates.
- Sherif, Muzafer (1937), “An experimental approach to the study of attitudes,” *Sociometry*, 1, 90–98.
- Shiota, Michelle, Dacher Keltner, and Amanda Mossman (2007), “The nature of awe: elicitors, appraisals, and effects on self-concept,” *Cognition and Emotion*, 21, 944-963.
- Sperber, Dan (1996), *Explaining culture: A naturalistic approach*, Oxford, England: Blackwell.
- Stephen, Andrew T. and Donald R. Lehmann (2009), “Why Do People Spread Word-of-Mouth? Effects of Recipient and Relationship Characteristics on Transmission Behaviors,” INSEAD working paper.
- Thibaut, John W. and Harold H. Kelley (1959), *The Social Psychology of Groups*, New York: Wiley.
- Van den Bulte, Christophe and Stefan Wuyts (2007), *Social Networks and Marketing*, Cambridge, MA: Marketing Science Institute.
- Van den Bulte, Christophe and Yogesh V. Joshi (2007), “New Product Diffusion with Influentials and Imitators,” *Marketing Science*, 26 (3), 400-21.
- Watts, Duncan J. (2002), “A simple model of global cascades on random networks,” *Proceedings of the National Academy of Sciences*, 99, 5766-5771.

- Watts, Duncan J. (2004), "The 'New' Science of Networks," *Annual Review of Sociology*, 30, 243–70.
- Watts, Duncan J. and Peter Sheridan Dodds (2007), "Influentials, networks, and public opinion formation Journal of Consumer Research," *Journal of Consumer Research*, 34, 441-458.
- Wojnicki, Andrea C. and Dave Godes (2008), "Word-of-Mouth as Self-Enhancement," University of Toronto Working Paper.

FIGURE 1

HOMEPAGE LOCATION CLASSIFICATIONS. PORTIONS WITH “X”’S THROUGH THEM ALWAYS FEATURED AP AND REUTERS NEWS STORIES, VIDEOS, BLOGS, OR ADVERTISEMENTS RATHER THAN ARTICLES BY *TIMES* REPORTERS

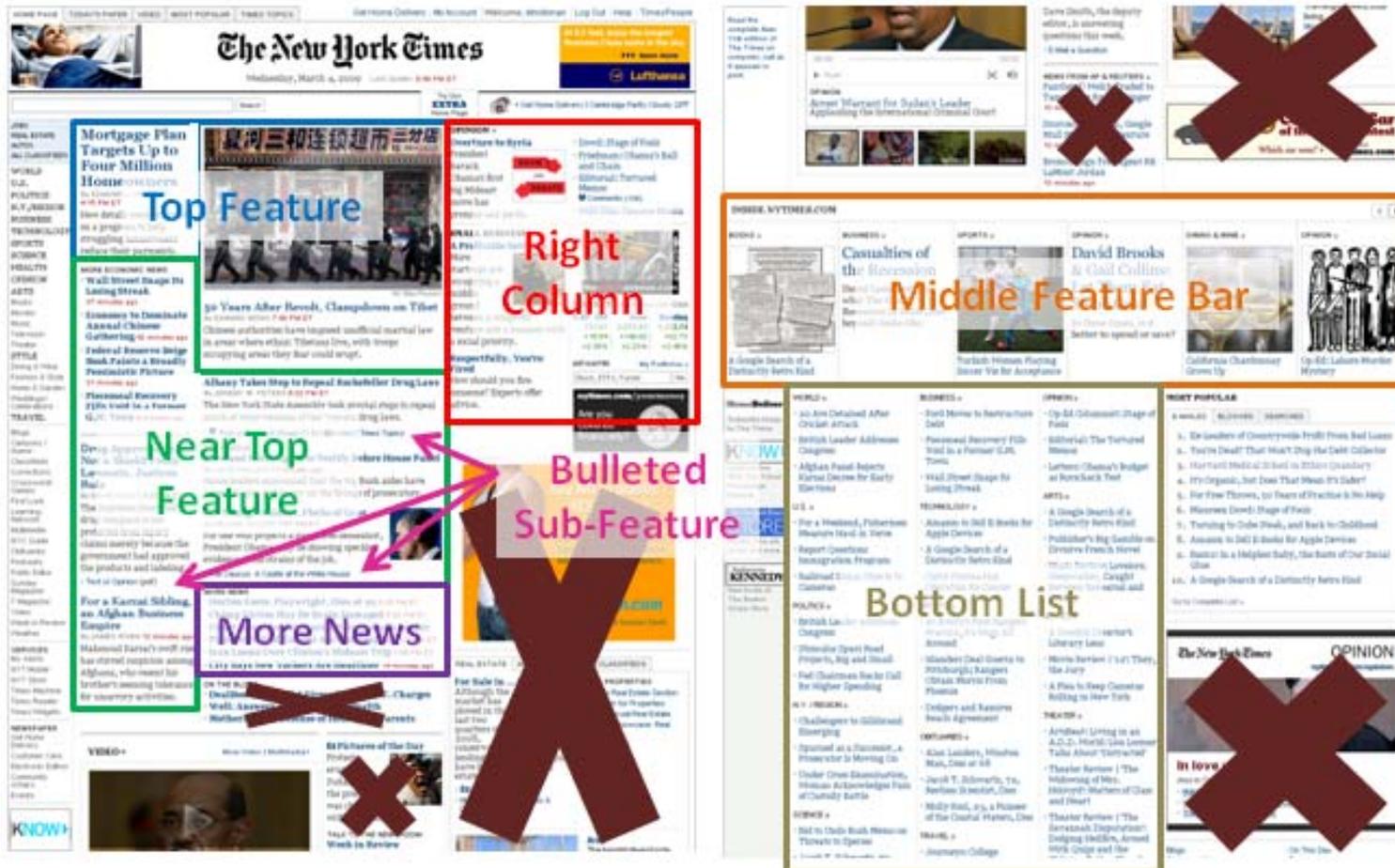


FIGURE 2

PROBABILITY AN ARTICLE MAKES THE *NEW YORK TIMES*' MOST EMAILED LIST AS A FUNCTION OF PSYCHOLOGICAL CHARACTERISTICS OF ITS CONTENT AND HOW PROMINENTLY IT IS FEATURED

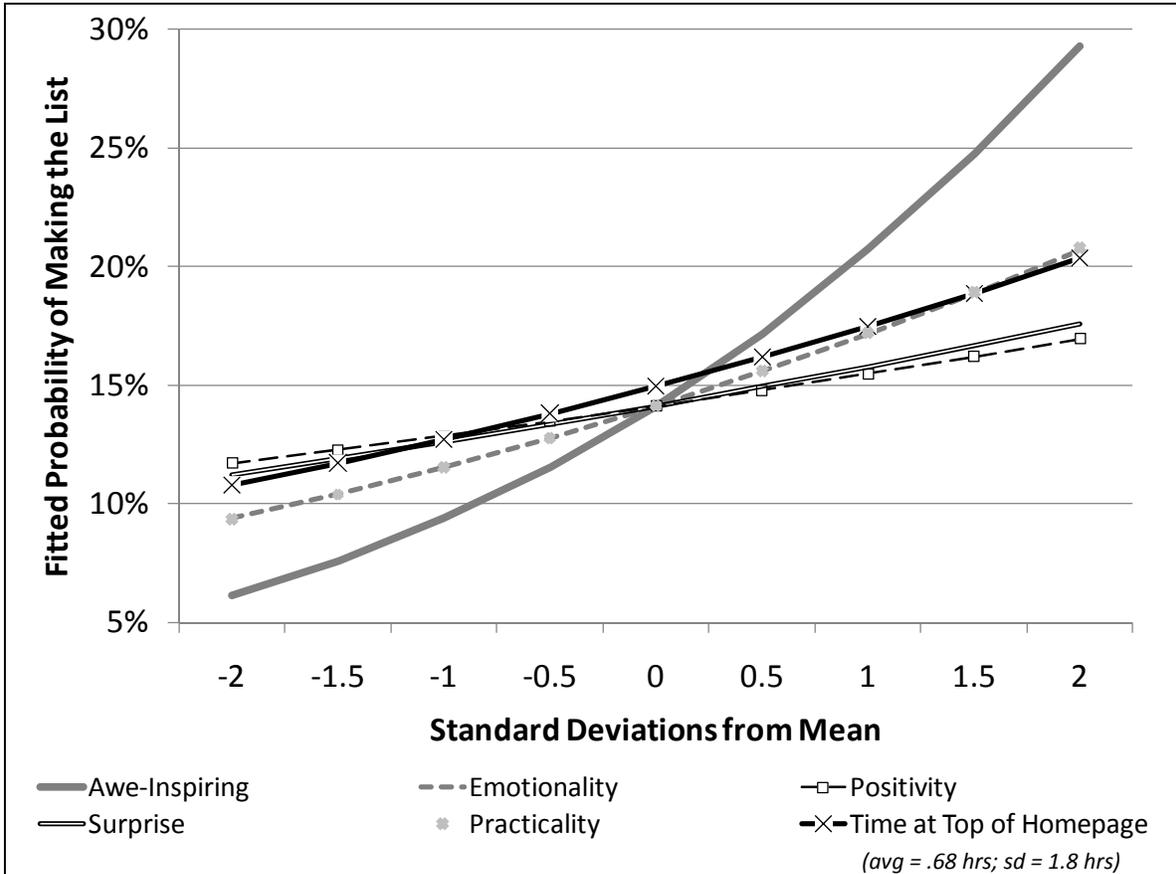


TABLE 1
PREDICTOR VARIABLE SUMMARY STATISTICS

		Mean	Std. Dev.
Primary Predictor	Awe-Inspiring*	1.73	0.74
Variables	Practicality*	1.76	1.19
	Surprising*	2.56	0.95
	Emotionality*	7.43%	1.91%
	Positivity*	0.99%	1.85%
Other Control	Wordcount	1,033.36	740.58
Variables	Complexity*	11.10	1.53
	Author Fame	9.14	2.53
	Author Female	0.29	0.45
	Author Male	0.66	0.47

*Note that these summary statistics pertain to the variable in question prior to standardization.

TABLE 2
CORRELATIONS BETWEEN PREDICTOR VARIABLES

	Word Count x 10⁻³	Complexity	Author Fame	Author Female	Emotionality	Positivity	Surprise	Awe-Inspiring
Word Count x 10⁻³	1.00							
Complexity	-0.06*	1.00						
Author Fame	-0.05*	0.02	1.00					
Author Female	-0.01	-0.03*	0.00	1.00				
Emotionality	0.07*	0.05*	-0.10*	-0.07*	1.00			
Positivity	0.06*	-0.05*	-0.04*	0.06*	0.05*	1.00		
Surprise	0.01	0.06*	0.04*	0.11*	-0.15*	-0.08*	1.00	
Awe-Inspiring	0.12*	-0.03	0.01	0.08*	0.11*	0.13*	0.31*	1.00
Practical	-0.02	-0.01	-0.04*	0.06*	0.08*	0.04*	-0.07*	0.18*

*Significant at 5% level.

TABLE 3
ARTICLE'S LIKELIHOOD OF MAKING THE *NEW YORK TIMES*' MOST E-MAILED LIST BASED ON PSYCHOLOGICAL CHARACTERISTICS OF ITS CONTENT, AS WELL AS VARIOUS CONTROL VARIABLES

	(1)	(2)	(3)	(4)
Awe-Inspiring	0.57*** (0.04)	0.49*** (0.04)	0.46*** (0.05)	0.40*** (0.06)
Other Psychological Variables				
Emotionality	-	0.25*** (0.03)	0.23*** (0.04)	0.11*** (0.04)
Practicality	-	0.08** (0.04)	0.23*** (0.05)	0.07 (0.05)
Surprise	-	0.12*** (0.05)	0.13** (0.05)	0.21*** (0.06)
Positivity	-	0.08*** (0.03)	0.11*** (0.04)	0.09** (0.04)
Homepage Location Control Variables				
Top Feature	-	-	0.10*** (0.02)	0.09*** (0.02)
Near Top Feature	-	-	0.09*** (0.01)	0.08*** (0.01)
Right Column	-	-	0.14*** (0.01)	0.09*** (0.01)
Middle Feature Bar	-	-	0.06*** (0.00)	0.05*** (0.00)
Bulleted Sub-Feature	-	-	0.04*** (0.01)	0.05*** (0.01)
More News	-	-	-0.01 (0.01)	0.05*** (0.01)
Bottom List x 10	-	-	0.05*** (0.01)	0.07*** (0.02)
Other Control Variables				
Word Count x 10⁻³	-	-	0.49*** (0.06)	0.66*** (0.06)
Complexity	-	-	0.02 (0.04)	0.03 (0.04)
First Author Fame	-	-	0.17*** (0.02)	0.15*** (0.02)
Female First Author	-	-	0.30*** (0.07)	0.28*** (0.08)
Uncredited	-	-	0.41* (0.23)	-0.63** (0.28)
Newspaper Location & Web Timing Controls	No	No	Yes	Yes
Article Section Dummies (arts, books, etc.)	No	No	No	Yes
Observations	7,710	7,710	7,710	7,710
McFadden's R²	0.05	0.06	0.24	0.31
Log likelihood	-3,677.80	-3,633.41	-2,930.65	-2,653.45
Likelihood-ratio chi² (A nested in B)		88.77***	1,405.52***	554.40***

*Significant at 10% level. **Significant at 5% level. ***Significant at the 1% level.

Relative effect sizes of coded variables should be interpreted with care, as these variables are necessarily proxies for underlying constructs rather than exact measures of those constructs. Of greater interest is the large relative estimated effect of each of these proxy variables on an article's likelihood of making the most e-mailed list compared to the cleanly measured effect of external drivers of attention (e.g., time spent on various positions on the *Times* homepage).

APPENDIX

Descriptions provided to coders.

Practical Utility. Articles vary in how much practical utility they have. Some contain useful information that leads the reader to modify their behavior. For example, reading an article suggesting certain vegetables are good for you might cause a reader to eat more of those vegetables. Similarly, an article talking about a new Personal Digital Assistant may influence what the reader buys. An article about travel might influence where the reader goes on vacation. Importantly, practical utility focuses on the here and soon. An article about cosmology, for example, might be very eye-opening, but if it doesn't affect what the reader does soon, that should not be coded as having high practical utility. Also focus on utility for the reader. If it won't potentially lead them to behave differently, it shouldn't be coded as having practical utility.

Awe-Inspiring. Articles vary in how much they inspire awe. Such articles teach people things or share a perspective that opens people's minds. This knowledge may not change how people live their everyday life, and thus may have little practical value, but it is more about knowledge for knowledge's sake. For example, reading an article about how the human mind works, or the structure of the universe is unlikely to change what people eat or what they do this weekend, but it will open up their mind to new possibilities and insights. Articles that are high on this dimension may evoke feelings of awe and wonder, and even articles that are mildly high on this dimension may just shift how the readers sees things, even if it doesn't really directly have practical utility.

Surprising. When thinking about sending an article to someone else, people may consider whether the information in the article will be "surprising" to the person receiving it. This can happen in two ways. One possibility is that the recipient will find it shocking. If an article suggested that milk killed brain cells, that would be very surprising. Another possibility is that the recipient will have a very low likelihood of being aware of the information contained in the article already. If milk killed brain cells, but everyone had already heard that news, then it would not be very surprising. Similarly, if the Yankees win the World Series that would be interesting, but given that most people would already know about it, it would not be very surprising.

Alternate Measure of Awe. Articles vary in how much they inspire awe. Awe is the emotion of self-transcendence, a feeling of admiration and elevation in the face of something greater than the self. It involves the opening or broadening of the mind and an experience of wow that makes you stop and think. Seeing the Grand Canyon, standing in front of a beautiful piece of art, hearing a grand theory, or listening to a beautiful symphony may all inspire awe. So may the revelation of something profound and important in something you may have once seen as ordinary or routine or seeing a causal connection between important things and seemingly remote causes.

TABLE A1
 HOMEPAGE LOCATION ARTICLE SUMMARY STATISTICS

	% of Articles That Ever Occupy This Location	For Articles that Ever Occupy Location:		
		% That Make List	Mean Hrs	Hrs Std. Dev.
Top Feature	26%	32%	2.63	2.97
Near Top Feature	31%	31%	5.00	5.09
Right Column	21%	31%	3.80	5.07
Middle Feature Bar	23%	32%	11.52	11.62
Bulleted Sub-Feature	34%	25%	3.62	4.11
More News	28%	24%	3.65	4.16
Bottom List	81%	20%	23.00	28.25

Note: The average article in our data set appeared somewhere on the *Times* homepage for a total of 27 hours (standard deviation = 30 hours),

TABLE A2
PHYSICAL NEWSPAPER ARTICLE LOCATION SUMMARY STATISTICS

	% of Articles That Ever Occupy This Location	For Articles that Ever Occupy This Location:		
		% That Make List	Mean Pg #	Mean Pg # for Articles that Make List
Section A	40%	25%	15.52	11.96
Section B	15%	11%	6.35	3.71
Section C	9%	16%	4.05	3.78
Section D	7%	19%	3.01	1.65
Section E	3%	22%	4.74	4.72
Section F	2%	42%	3.28	2.02
Other Section	13%	24%	9.47	13.64
Never in Paper	11%	10%	-	-